International Data Base
Population Estimates and Projections Methodology

June 2009

Overview

The U.S. Census Bureau has been preparing estimates and projections of the populations of foreign countries since the 1960s. In the 1980s the Census Bureau released its first comprehensive set of estimates and projections for over 200 countries and areas of the world. Since then, the Census Bureau has routinely updated estimates and projections for countries as new data have become available. Estimates and projections for countries, as well as for regions and the world, are made available to the public, through the Census Bureau's International Data Base (IDB), located at Web site http://www.census.gov/ipc/www/idb/.

The Census Bureau's IDB estimates and projections have several distinguishing features. For countries and areas recognized by the U.S. Department of State and which have populations of 5,000 or more, population size and components of change are provided for each calendar year beyond the initial or base year, through 2050.[1] Within this time series, sex ratios, population, and mortality measures are developed for single ages through age 100-plus. As a result of single-year age and calendar-year accounting, IDB data capture the timing and demographic impact of important events such as wars, famine, and natural disasters, with a precision exceeding that of other online resources for international demographic data.

The estimation and projection process involves data collection, data evaluation, parameter[2] estimation, making assumptions about future change, and final projection of the population for each country. The Census Bureau begins the process by collecting demographic data from censuses, surveys, vital registration, and administrative records from a variety of sources. Available data are evaluated, with particular attention to internal and temporal consistency.

The U.S. Census Bureau strives to base the population estimates and projections for each country on a modified *de facto* population universe whenever possible. A strict interpretation of the *de facto* population includes all persons who are physically present in the country at the reference date, whether or not they are usual and/or legal residents. In contrast, the *de jure* population consists of all usual residents, whether or not they are present at the reference time. Our estimates and projections exclude foreign military

---

[1]The Census Bureau provides estimates and projections of West Bank and Gaza Strip areas, in addition to the countries and areas recognized by the U.S. Department of State.

[2] A parameter is a population characteristic, such as age, sex, and fertility, mortality, and migration levels and patterns.

populations, tourists and others visiting for short periods of time from the country's *de facto* population. It is not always easy to adjust the data to represent this modified *de facto* population universe. After appropriate adjustments are made, the data are then used to estimate population by single years of age, as well as to estimate the fertility, mortality and migration parameters needed for population projection.

Estimation and projection procedures make use of a variety of demographic techniques and incorporate assumptions formed by consulting the social science and health science literature. In addition to using demographic data, Census Bureau demographers consider information on public health efforts, socio-political circumstances, and historical events such as natural disasters and civil conflict in preparing the assumptions feeding into population projections. The best available estimates and carefully derived assumptions are then placed as inputs into the Census Bureau's Rural/Urban Projection (RUP) program.[3] Finally, the RUP program processes the inputs and generates cohort component projections up through 2050 for each country. Regional and world populations are obtained by projecting each country's population separately and then combining the results to derive aggregated totals.

The use of RUP applies to all countries except the United States. The estimates and projections for the United States provided in the IDB are prepared using data and procedures that differ slightly from those used for all other countries. The population estimates, which span 1950 to the calendar year prior to the current year, [4] are developed using decennial census counts and information about vital statistics and international migration. The IDB estimates for 1950 to 1999 are based on the resident population plus the armed forces overseas. [5] The IDB estimates for 2000 to the last calendar year are based on the Census 2000 resident population only.

The IDB population projections, which span the current calendar year to 2050, are projections of the resident population that are based on Census 2000 and projected forward using historical trends in vital statistics and international migration. The projections are produced by applying the cohort-component method to assumptions about fertility, mortality, and international migration. There is a break in the time series in the current calendar year because the population estimates and the population projections are not entirely consistent.

For years in between the most recent census and the present, the recommended source for U.S. population estimates is the Census Bureau's Population Estimates Program, which produces the official estimates of the U.S. population for the country. These estimates, as

---

[3] For more information on the U.S. Census Bureau's Rural/Urban Projection program, see http://www.census.gov/ipc/www/rup/index.html
[4] New projections replace estimates for the U.S. with the June IDB update each year.
[5] These estimates are from the Census Bureau's Population Estimates Program. Since 1980, the Estimates Program has been producing separate estimates for the two population universes of (1) the resident population and (2) the armed forces overseas population. Using estimates of these two universes, the population Estimates Program had produced estimates for a third population universe, the "resident population plus armed forces overseas" population. Prior to 1980, the Estimates Program did not separate out the residents from the armed forces overseas population.

well as a more complete description of the methodology, can be found at
http://www.census.gov/popest/estimates.html. Information about the population
projections is available at
http://www.census.gov/population/www/projections/index.html.

Although the Census Bureau applies methodologically sound demographic techniques to
evaluate and estimate population and its components of change, the accuracy of the
estimates and projections in the IDB is limited. First, the accuracy of parameter estimates
and assumptions can be compromised due to deficiencies in their source data. Second,
regardless of the accuracy achieved in estimating the current demographic situation in
countries, assumptions about the future can turn out to be highly inaccurate, since some
population dynamics are influenced by events that are unpredictable – natural disasters
and the scale of their impact on a population, for example, cannot be foreseen.

The Cohort-Component Method

A cohort-component population projection follows each group of people of the same age
and sex throughout its lifetime, exposing it to assumed age- and sex-specific mortality,
fertility, and migration.

An initial or base year population, disaggregated by age and sex, is exposed to estimated
age-sex-specific chances of dying as determined by estimated and projected mortality
levels and age patterns. Once deaths are estimated, they are subtracted from each age,
yielding the next older age in the subsequent time period. Although 5-year age groups
can be used in implementing the cohort-component method, the U.S. Census Bureau uses
single years of age in developing the estimates and projections provided in the IDB.
Fertility rates are projected and applied to the female population of childbearing age to
estimate the number of births every year. Each cohort of children born is also followed
through time and survivors are calculated after exposure to mortality. Finally, the cohort-
component method takes into account immigrants who are incorporated into the
population and emigrants who leave the population. Net migrants are added to, or
subtracted from, the population at each specific age. The whole procedure is repeated for
each year of the projection period, resulting in the projected population by age and sex,
through 2050.

Making a cohort-component population projection is a multi-step process. Its complexity
lies not in the calculations required – computer programs handle this quickly and
painlessly – but, rather, the derivation of projection inputs: parameter estimates and
assumptions about future change. An essential step in deriving these inputs is the
development of estimates that accurately represent the current *de facto* population, as
well as past and present trends in fertility, mortality, and migration. The development of
estimates involves gathering available demographic data, assessing their quality,
adjusting them as necessary using demographic techniques, and reviewing their
comparability among countries. These estimates, in the context of the projection process,

are referred to as "base estimates" or "base data." Once base estimates are developed, they can be used to make reasonable and consistent assumptions about the future course of fertility, mortality, and international migration. Both the base estimates themselves, and the assumptions about the future which are formulated by the base estimates, comprise the projection inputs which, when integrated and subjected to mathematical calculations embedded in RUP, generate population projections from the base year through 2050.

The following sections describe in more detail the processes of developing base estimates and, thereafter, generating population projections. Also provided is information on data sources. It is important to note that RUP allows the entry of population sex and age structure estimates only for the initial or base year. Thereafter, later base estimates for the population are used in evaluating the RUP generated population but are not entered into the program.

Population: data and base estimates

The base population is most often derived from census data if available and of acceptable quality. Another source of the base population is official population estimates produced by the national statistical office. Typically, countries which produce good population estimates also have vital registration systems with complete coverage of births and deaths and administrative data that captures international migration. For a limited number of countries, neither an acceptable census nor official estimates are available and a stable population is generated.

Census enumerations are not perfect: reported data on the population may be affected by several factors. The total population may be undercounted due to the inability of census enumerators to cover the country adequately. Furthermore, the age and sex structure may be affected by age misreporting, by underenumeration of people of certain ages, and in some cases by duplications of people of certain ages. Before using the census counts as the base for projection, they must be evaluated for errors and adjusted when necessary.

In evaluating census age-sex structures for base population estimates, special attention is given to possible underenumeration of the youngest age groups, 0-4 years and 5-9 years. Underenumeration of the youngest age groups is common in census data of most countries, although the degree of the problem varies. Errors in these ages may have significant impacts on the cohort-component projection. Suppose, for example, that children ages 0-4 were underestimated in the base population. In the projection, not only would the surviving cohorts of these children be smaller than they should be, but when the projected female cohorts reach reproductive ages, their number of births would also be underestimated.

Various methods are applied to detect errors in the census population before it is accepted for use in developing a base population. In evaluating the census population, the errors

that the Census Bureau aims to detect are those affecting the total population and its age and sex distribution. While the true values of a country's population and its population by age can rarely be known, these values can often be approximated. Demographic methods can be used to measure net census error; that is, the difference between what is considered to be the correct value and the value measured by the census.

The main approach for evaluating a census is referred to as "demographic analysis." This approach essentially entails calculating demographic indices using data from the census under study, and if available, results from other censuses and additional demographic data sources. Using data from the census under study, crude and approximate measures can be made to check for errors in the count of persons. For example, the Census Bureau determines if the census covered all of the country's geography. For some countries, especially those engaged in border disputes or wars, geographic sub-entities are left out of the census altogether. The Census Bureau identifies alternate data sources to measure the population of these omitted areas.

Applying more rigorous demographic techniques, the Census Bureau further utilizes the census results to detect errors in population distributions by sex and age. These techniques entail the analysis of distributions or ratios, including analysis of digit preference, age ratios, and sex ratios. [6] In some cases, analyses are conducted by comparing the census age distribution with an age distribution that would have resulted from constant fertility and mortality over a fairly long period of time and which has not experienced international migration—"stable" population. Through this comparison, age and sex distribution irregularities are detected.

Evaluating a census through demographic analysis is also undertaken using data from previous censuses of the same population. Using data from a previous census, an

---

[6] Indices frequently used by the Census Bureau for detecting digit preference in base population data for the IDB include those developed by Myers (1940), Whipple (U.S. Bureau of the Census, 1971), and Bachi (1951, 1953). These indices provide an overall idea of the extent of age misreporting as well as the preference for certain ending age digits. Age ratio analysis (United Nations, 1952) using population data for 5-year age groups is used to detect age misreporting in populations where fertility has not fluctuated greatly during the past and where international migration has not been significant. Calculations entail dividing the population in a specific 5-year age group by the average population of the two adjacent 5-year age groups, times 100. The larger the fluctuations of these ratios and the larger their departure from 100, the greater the probability of errors in the data. Sex ratios are simply calculated by dividing the male population by the female population in a given age or age group, times 100. For most middle age groups, and depending on the level of sex-specific migration in a population, the larger the departure from 100, the larger the possibility for errors in the data. For the youngest ages, 0 or 0-4, the sex ratio tends to range between 103 and 105 due to biological factors which result in an excess of male births. For some countries, sex ratios for births occurring in the latter half of the 20th century have dramatically increased, to levels above 112 for example, as a result of sex-selective abortions where some parents have a preference for male children over female children. For the oldest age groups, the number of females tends to exceed males; however there are exceptions to this trend as well, largely in societies that favor males. Joint age- and sex-indices -- the sex-ratio score, the age-ratio score, and the age-sex accuracy index -- developed by the United Nations (1952 and 1955) are used as summary measures of the age and sex ratios.

expected census count of population can be calculated and compared with the census under consideration for base population estimation. More specifically, the previous census total and an assumed rate of intercensal growth can be used to project an alternate population for the base population year; and this alternate population—considered the "expected population"—can be compared with the census under study as a measure of census accuracy. The assumed rate of intercensal growth is calculated using a total population from an additional previous census, or through a more elaborate estimation strategy. The more elaborate estimation strategy entails calculating intercensal growth separately for each of the components of population change—fertility, mortality, and international migration.[7] The level of sophistication applied in estimating intercensal growth varies depending on data availability.

Given that errors in the under age-10 census population have been frequently found, demographic analysis oftentimes focuses on examining this age group. The completeness of enumeration for this young population is often evaluated by checking for consistency between its enumerated size and estimated levels of fertility and mortality during the 10-year period prior to the census date, as children under age 10 at the census date represent the survivors of births during the 10-year period prior to the census date. More specifically, consistency is checked by first estimating the number of females of reproductive age during the two 5-year periods prior to the census date, using the census data themselves and mortality information from another demographic data source. Based on this estimated female population and fertility and mortality levels measured by other demographic data, the male and female births during the two 5-year periods prior to the census date are then estimated. The two estimated 5-year birth cohorts are then projected to the census date, when they will have aged to 0-4 and 5-9 years old. Finally, the projected populations ages 0-4 and 5-9 are compared to the enumerated census population ages 0-4 and 5-9, in order to gauge the census enumeration accuracy.

The Census Bureau also uses a later census to evaluate the base population's under-10 and over-10 estimates. This comparison is performed following the estimation of inputs for the cohort component projection (which will be more fully elaborated in the sections on mortality, fertility, and net international migration) and the subsequent projection of the base population using the components of change.

For some countries, censuses can be evaluated using results from "Post Enumeration Surveys" (PES). The volume of information available from a PES varies. Ideally two different types of information will be available for a given census: (1) information from a sample survey, taken shortly after the census, wherein residents who responded to census questionnaires are re-interviewed and, less commonly, (2) information from a post-censal matching study that consists of the re-enumeration of an independently-selected probability sample of the target census population. The Census Bureau will

---

[7] This approach involves evaluating the quality of data formulating assumptions about intercensal change mortality, fertility, and international migration. Data sources for these assumptions include surveys, vital registration data, administrative records, and, in some cases, estimates of rates from country statistical offices or international statistical organizations.

consult the results of both types of PES operations, if available, as part of the process of evaluating the census results that will be used to derive the base population.

In addition to demographic analysis and review of PES results, evaluations of the census population or subpopulations can be made in a rough manner, via comparing them to administrative statistics or survey measures. These aggregate level analyses are often executed using Demographic and Health Survey (DHS) data, available for many countries.[8]

Once errors are detected, techniques are applied to make needed adjustments. Younger age groups found to be underenumerated are adjusted upward using the results of one of the various aforementioned demographic analysis techniques. If warranted, the age distribution for the older age groups might be adjusted using a smoothing technique.[9]

For some countries, age data from high quality national-level surveys can be used to adjust problematic census age data. Data from some national-level surveys can better capture the age distribution of country populations since more resources are generally devoted to structuring survey questions and training interviewers for nationally representative surveys, than can be devoted to most censuses. Country DHS surveys collect age data which the Census Bureau sometimes uses in adjusting a census age structure. Oftentimes a combination of such survey data and census data is used. For example, the census totals by sex might be disaggregated by 5-year age groups, using survey data. Despite the relatively greater accuracy surveys have in capturing the under-10 population; there still tends to be some undercount in this group, so the Census Bureau examines the data using the same types of techniques used to examine census data.

Once any needed adjustments are performed for the base population, an extra step of "moving" the base population to midyear[10] is performed, using information about fertility, mortality, and migration during the short period between the census reference date and midyear.

After the base population is moved to midyear, the oldest, open-ended age group usually needs to be extended to represent age groups up through 100 years and over.[11] Because many country censuses do not tabulate and publish age data for these older age groups and because age misreporting may be especially problematic in older populations, the Census Bureau uses models that take into account assumptions about mortality and

---

[8] The Demographic and Health Surveys Programme started in 1984. It was designed to gather demographic data in countries which did not have vital registration systems that adequately captured births and deaths. Under DHS auspices, surveys have been carried out in more than 70 countries in Africa, Asia, Latin America, the Caribbean, and parts of Europe.

[9] Arriaga and Associates (1994: 39-42) provide a comprehensive description of demographic techniques in *Population Analysis with Microcomputers*.

[10] The Census Bureau traditionally has used July 1 as the midyear date for its population estimates and projections.

[11] Extension of the base population to 100 years and over can be performed before or after it is moved to midyear.

population size for various enumerated age groups 55 and older. Currently, the adjustments applied draw from stationary and stable population models.

Although no country's base population is entirely unique, there are some typical patterns. In developing base populations, the Census Bureau knows that, based on empirical evidence compiled over many years of research, certain types of countries will assume different types of patterns, and they use this knowledge as part of the process of assessing the quality of base populations' overall age-sex structure.

The age-sex structures graphed are often referred to as "population pyramids." For more developed countries that have experienced lower fertility and higher life expectancies in the latter half of the 20th century, the term "population pyramid" is somewhat of a misnomer. A graphed depiction of the age-sex structure for these more developed countries tends to deviate from a classic pyramid shape. Specifically the base of the "pyramid" tends to narrow, as a reflection of lower birth rates. This type of age-sex structure is illustrated in Finland's census population "pyramid" for 1990, below.
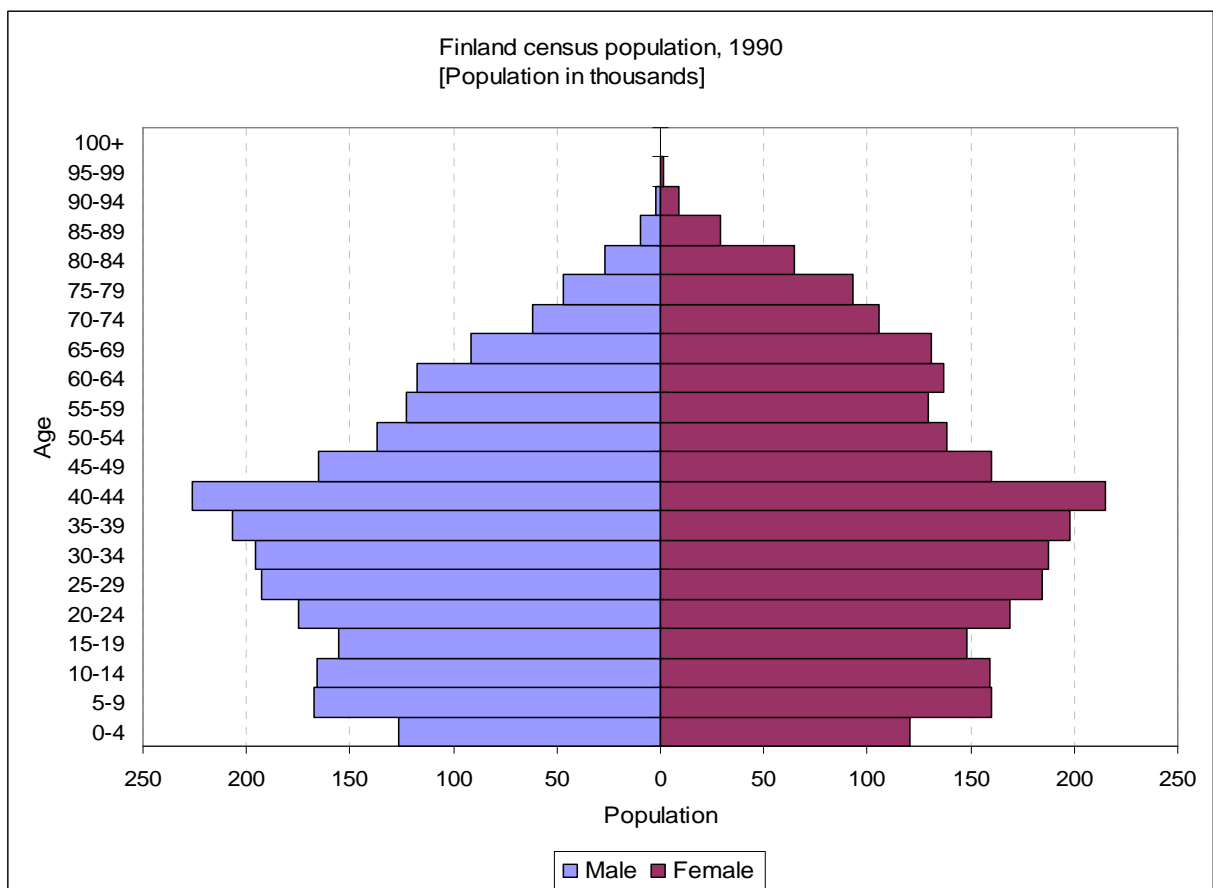
Source: Census 1990 of Finland as published in the *Demographic Yearbook, Historical Supplement 1948-1997* (United Nations, 2000).

For some country populations, and especially those dated relatively further in the past, more classically shaped pyramids represent the age-sex structure. The 1979 census population for Afghanistan, for example, has a classical pyramid shape.

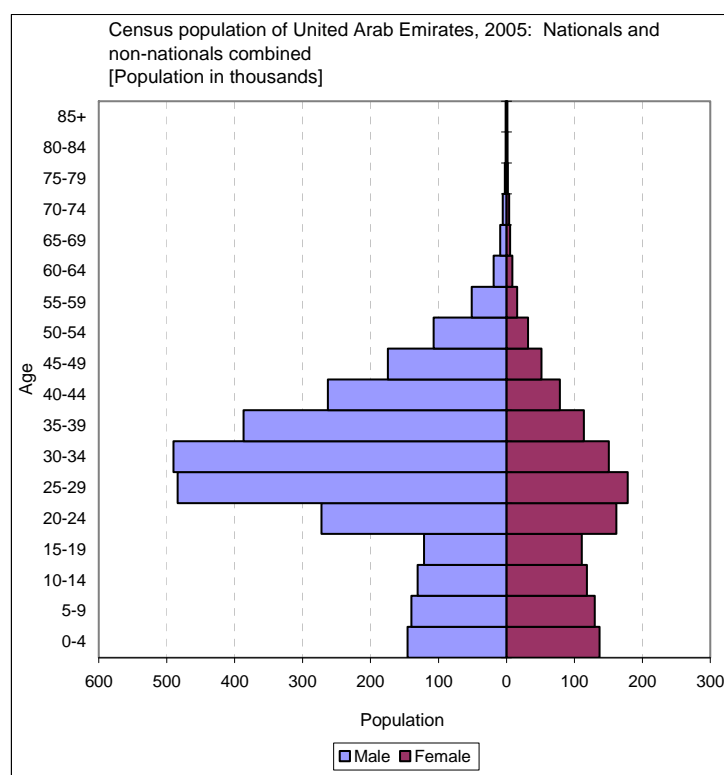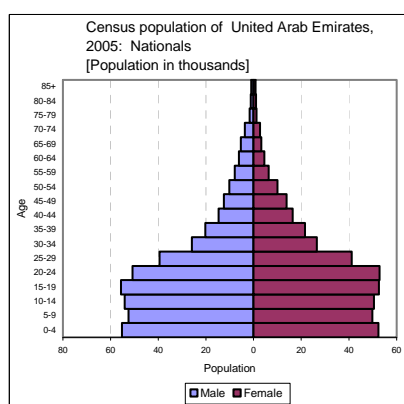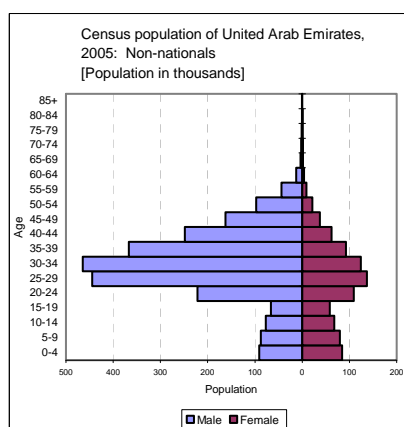Afghanistan census population, 1979
[Population in thousands]
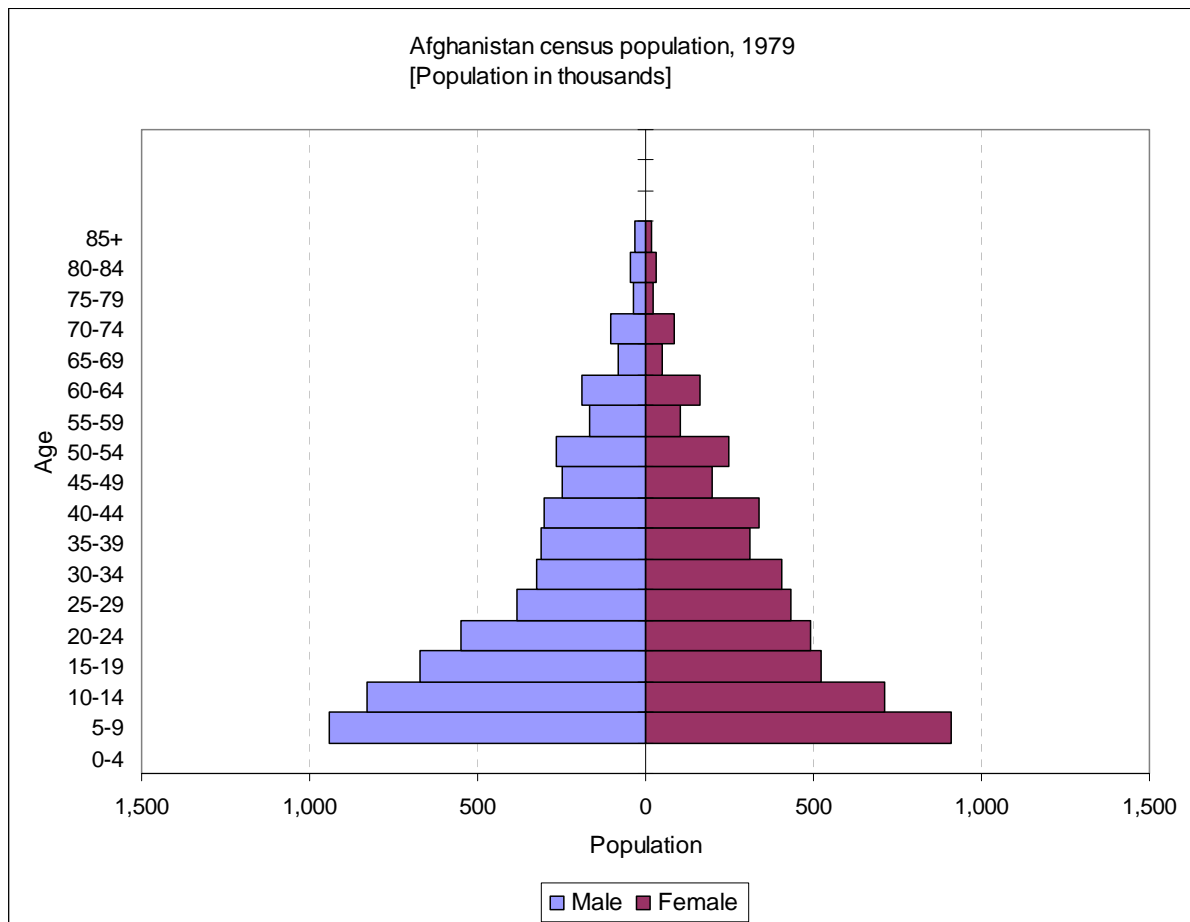
Source: Census 1979 of Afghanistan as published in the *Demographic Yearbook, Historical Supplement 1948-1997* (United Nations, 2000).

Assessments of census coverage errors can go beyond looking for the common problem of under-10 underenumeration and age misreporting at older ages. Sometimes coverage errors occur in other age groups as well, depending on the political, economic, and social events in a particular country. An apparent dearth in the number of men of military age, for example, can occur when a particular population subgroup is at odds with the national government. Similarly, underenumeration of the working age population can occur when part of this age group consists of unauthorized immigrants, who may deliberately evade census enumeration for fear of apprehension. On the other hand, excesses in some

population groups can be found, depending on their relationships with the national government. Adjusting for these coverage errors usually entails evaluation and adjustment based on comparison with prior censuses and population change in the interim period, taking into account evidence from special reports or articles about the group in question.

The Census Bureau also has to look beyond census numbers in order to decipher problems in data collection from real dearths—or in some cases, excesses—in a population for certain age groups and for one gender or the other. This involves knowing the history of a country and examining the extent to which past events have affected the population structure. Some countries with a history of civil strife enumerate relatively small numbers in the male cohorts who had been of military-age at the time of civil unrest. These population losses, furthermore, remain in the population structure as time progresses, and the Census Bureau must calculate the timing of the historical event and ensure that the markedly small population size of a particular cohort matches the date of a significant event. The effects of war in 1963 on males ages 20-24, might be apparent in a 2013 population, among males ages 70-74, for example. For other countries, special adjustments must be performed in order to ensure excesses of certain age groups are accurately represented. In many Persian Gulf States, for example, a large proportion of the labor force consists of international migrants, who are predominantly male. Because of this particular feature in Persian Gulf countries, the base population age-sex structures for these countries necessarily include a large male-dominant working-age population.

Below is the pyramid representing the 2005 preliminary census population for the United Arab Emirates (U.A.E.); as well as the two subgroups of this population: (1) U.A.E. nationals and (2) U.A.E. non-nationals who largely consist of international migrants who moved to the U.A.E. as labor migrants. The two subpopulations, combined, represent U.A.E.'s 2005 population in its entirety.

Census population of United Arab Emirates, 2005: Non-nationals [Population in thousands]

Census population of United Arab Emirates, 2005: Nationals and non-nationals combined [Population in thousands]

Census population of United Arab Emirates, 2005: Nationals [Population in thousands]

Sources: Preliminary Census 2005 of the United Arab Emirates, as published in the United Arab Emirates *Statistics Yearbook 2007* (United Arab Emirates).

Mortality: data and base estimates

Base estimates of mortality levels and age- and sex-specific patterns needed for cohort-component projections are derived in a variety of ways, depending on data available for each country. For many statistically more developed countries, base data on mortality are taken from vital registration systems. Official estimates of life expectancy at birth and mortality rates derived by national statistical offices are also sometimes used. For most

statistically less developed countries, however, the U.S. Census Bureau develops base mortality estimates using a combination of data sources: nationally-representative household surveys; to a lesser degree, censuses; and increasingly, vital registration data. Additionally, in order to capture unusual increases in mortality, or demographic "shocks," special tallies and research reports from various national and international organizations are used.

When a country has high quality death data, we often use the data directly in the cohort-component projection. In cases where country death data do not have a a high enough quality to allow for direct use in this way, the data are used as input in life table construction; and the age-specific central death rates generated from the life tables are then used as input in the cohort-component projection. Depending on the quality of the death data, the life table construction process varies in complexity. Use of death data follows an evaluation and adjustment of registered deaths or deaths reported in response to a census or survey question about deaths of household members during the past year; or indirect estimation from census or survey questions about child survival, orphanhood, and widowhood; or on intercensal population change.[12] Assessments of completeness and of errors affecting registered deaths from the United Nations Statistics Division's *Population and Vital Statistics Reports* and its annual *Demographic Yearbook*, and from the World Health Organization's Statistical Information System (WHOSIS) are consulted as a first step in deciding whether registered deaths from a specific country are likely to be usable without adjustment.
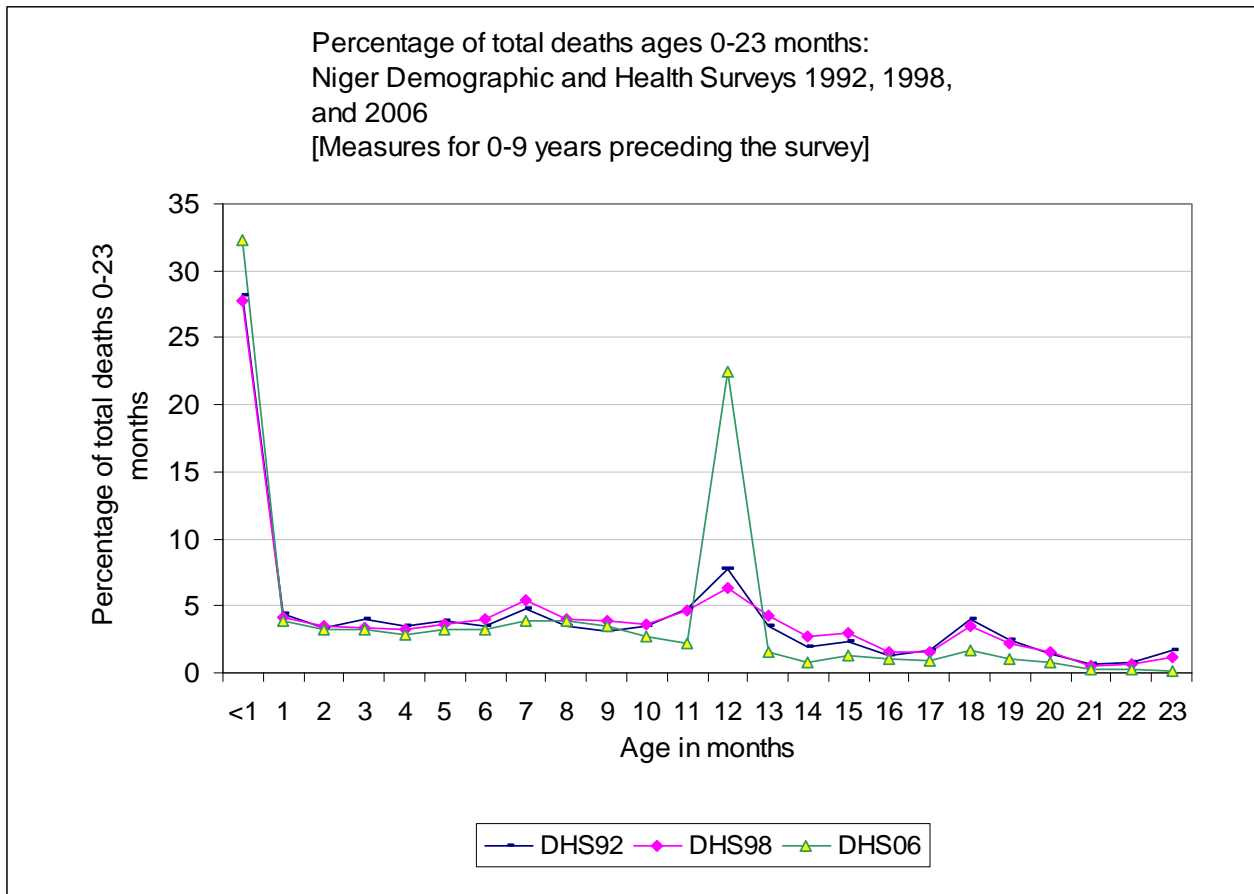
Mortality for all age groups is often estimated using information on infant and child survivorship collected in censuses and surveys as a guide to both level and age pattern. In the absence of either registered deaths or a census- or survey-based age distribution of deaths at adult ages, the distribution of deaths across age groups under age 5 provides evidence in the selection of a pattern of mortality at older ages. Both the Coale-Demeny regional model life tables (Coale and Demeny 1966) and the United Nations model life tables (United Nations 1982) are used to construct complete life tables consistent with estimated level and pattern of child mortality.

However, survey-based estimates of child mortality are not accepted without review. In a distribution of child mortality deaths by month, an examination of the magnitude of infant deaths in the 12th month category, compared with the number of deaths in the 0-11 month categories, and 13-plus month categories, may indicate heaping on the 12th month mark due to normal recall bias. Since a measure of infant mortality requires deaths from 0 through 11 months (the 12th month is not included in the measure) heaping of infant

---

[12] Methods used to evaluate and adjust both registered deaths and deaths during a 12-month period prior to a census or survey include (1) Brass' growth balance method and the Preston and Coale technique (United Nations 1983: ch. V) and (2) the general growth balance method (Hill 1987, Bhat 2002). Methods for estimating childhood mortality, including the Brass technique and the Palloni-Heligman reformulation, are described in United Nations (1983: ch.III) and United Nations (1990). The Johnson technique is presented in Arriaga and Associates (1994: 167-168). Methods based on surivorship of partents and siblings are described in United Nations (1983: ch. IV). Intercensal survival techniques are covered in United Nations (1983: ch. IX).

deaths on the 12th month results in an underestimate of infant mortality (and an over-estimate of mortality for children ages 1-4).

An example of age heaping found in child deaths data from Niger DHS surveys for 1992, 1996, and 2006 is shown here.

Percentage of total deaths ages 0-23 months:
Niger Demographic and Health Surveys 1992, 1998, and 2006
[Measures for 0-9 years preceding the survey]

Sources: *Niger Demographic and Health Survey 199*2 (Kourgueni et al., 1993), *Niger Demographic and Health Survey 1998* (Attama et. al, 1998), and *Niger Demographic and Health Survey 2006* (Institut National de la Statistique (INS) et Macro International Inc., 2007).

Where birth-history based estimates of infant and child mortality are available from multiple surveys for a country, those estimates are compared across measurement periods. For example, infant mortality rates (IMRs) measured for 0-4 years before a 1995 survey can be compared with IMRs measured 5-9 years before a 2000 survey as a check on consistency. For some countries, estimated under-5 mortality may be distorted by birth transference – misdating that selectively biases estimated child mortality levels (Sullivan 2008). In addition to internal data checks, part of any mortality data evaluation process may involve checking for consistency between adjusted or estimated mortality

levels and other demographic indicators for a country. For example, adjusted mortality levels can be evaluated by integrating them into the projections process and by checking the population projected using the mortality inputs against a reliable census or adjusted census population count.

Because country populations and the world population as a whole are aging, the Census Bureau now produces estimates and projections of mortality by single age for ages 0-99 and 100 and over for all countries. Although for some countries, this is a straightforward process, for others, which lack mortality data for these older age groups, models are used to estimate the age pattern of mortality at older ages. Even in countries where death data are available for older ages, the use of models for extending mortality up through the older ages is sometimes favored over relying solely on empirical data because empirical deaths from the smaller populations of older age groups are subject to greater age reporting error than those at younger ages. Empirical deaths from the smaller populations of older age groups are also more subject to year-to-year fluctuation; and are therefore less stable than are deaths from younger, larger age groups.

The models used to extend mortality to 100-plus begin with the assumption that underlying death rates of older age groups follow a smooth progression; and therefore, any fluctuations in empirical deaths rates are due to small numbers of events or errors in the data, which are masking the underlying smooth trends. The approach taken draws on research by Coale and Guo (1989) in extending the Coale-Demeny model life tables to age 100-plus using a logistic function.

Estimated age-sex patterns and time trends in mortality for some countries rely neither on standard demographic mortality models nor on the patterns found in similar countries due to special circumstances. Natural disasters and conflicts distort mortality levels and patterns, requiring additional modeling. Estimated numbers and the age-sex composition of deaths may be taken from reports by various national and international organizations, including the World Health Organization (WHO), the U.S. State Department, the U.S. Agency for International Development (USAID), and the U.S. Centers for Disease Control and Prevention (CDC). Additionally, some national statistical offices conduct their own death accounting and make the results available to the public. More often than not, these special mortality statistics are rough estimates and they do not include complete information on the distribution of deaths by age and sex. Furthermore, such accounts tend not to capture important patterns of immediate death increases and longer-term mortality impacts indirectly related to the unusual event.

Where information supporting the modeling of mortality associated with disasters and conflict is available, these data are incorporated into base mortality estimation. The impact of the 2004 Indian Ocean tsunami on the mortality rates in Indonesia, for example, was captured for Indonesia's base mortality estimates using special tallies of tsunami deaths combined with survey and other on-the-ground evidence indicating that the relative risk of tsunami deaths was higher for females than males. In other instances, where less information is available, simplifying assumptions are made in deciding how to distribute excess deaths by age and sex. Excess mortality resulting from the 2003

earthquake of Iran, for example, was calculated using a total death tally from the Iran State Statistical Centre and the age-sex distribution of Kerman ostan, the subnational area in which the earthquake occurred. The resulting shock to mortality levels, and resulting changes in life expectancy at birth are shown below.[13]

Expectation of Life at Birth for Iran: 1992-2009
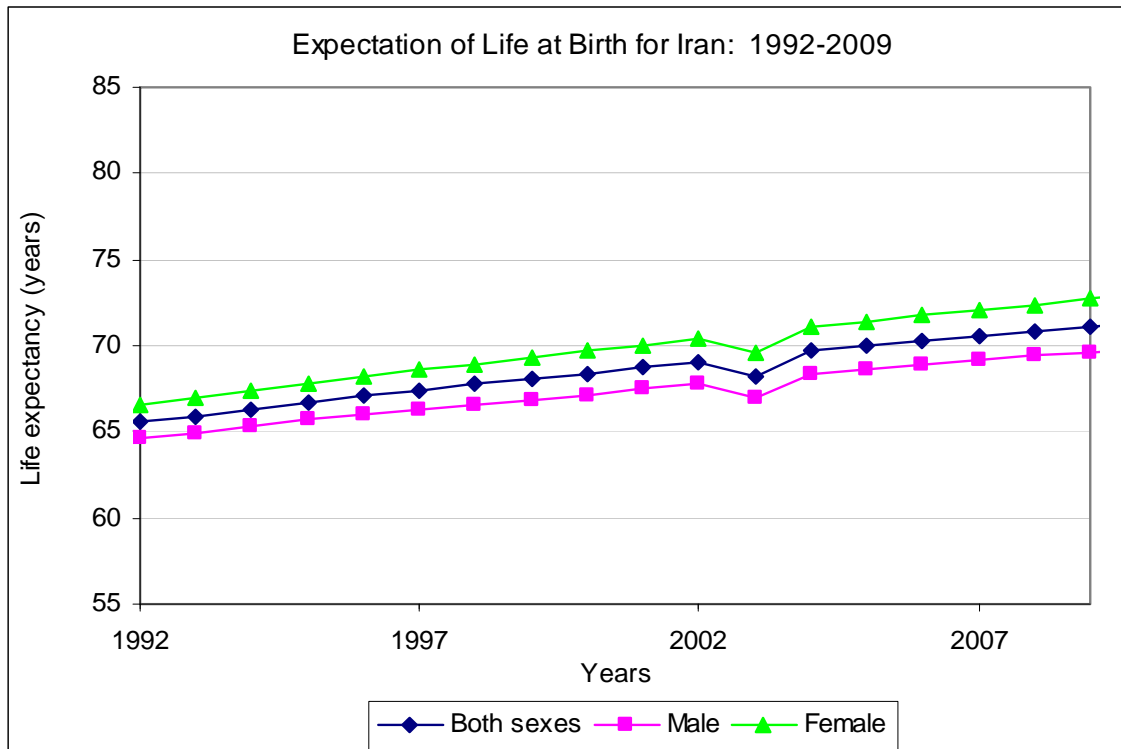
Source: Based on the U.S. Census Bureau's estimates and projections update for Iran, August, 2006.

Population Estimates and Projections Incorporating AIDS

The estimation and projection procedure for countries seriously affected by HIV/AIDS is somewhat complex. For these countries, the Census Bureau models mortality levels and trends under the hypothetical scenario of no epidemic, then adds estimated AIDS-related mortality based on measured HIV prevalence, ensuring that the "with-AIDS" mortality levels are consistent with empirical, population-based estimates. The starting point for the procedure is the Census Bureau's HIV/AIDS Surveillance Data Base.

The International Programs Center, Population Division, U.S. Census Bureau, maintains the HIV/AIDS Surveillance Data Base http://www.census.gov/ipc/www/hiv/. This database is a compilation of aggregate data from HIV seroprevalence and incidence studies in developing countries. Currently it contains over 100,000 records from various

---

[13]Life expectancy at birth for Iran measured in Iran's estimates and projections update, August, 2006.

publications, presentations, surveillance reports, and conference proceedings. HIV prevalence points taken from this database are the basis for projecting HIV prevalence and estimating AIDS mortality in countries that have generalized HIV/AIDS epidemics. A generalized epidemic is an epidemic that is widespread in the general population while a concentrated epidemic is concentrated in groups whose behaviors expose them to a high risk of infection.

The U.S. Census Bureau explicitly models AIDS-related mortality for those countries where adult HIV prevalence is consistently above one percent in the general population and transmission is mainly through heterosexual sex.

The impact of AIDS mortality is currently modeled explicitly in the estimates and projections for selected countries located in Asia, Latin America and the Caribbean, Sub-Saharan Africa, and Europe.

In 2004, a new application (RUPHIVAIDS) was developed at the Census Bureau to work with the Census Bureau's cohort component Rural-Urban Projection (RUP) program to model the impact of HIV/AIDS on the demography of a country. The RUPHIVAIDS model allows for competing risk of death from AIDS versus other causes.

RUPHIVAIDS uses estimates of HIV prevalence from the Estimation and Projection Package (EPP), an epidemiologically realistic model developed and used by the WHO and the Joint United Nations Programme on AIDS (UNAIDS). EPP produces a national "best fit" curve of adult HIV prevalence by year using sentinel surveillance data pertaining to pregnant women visiting antenatal clinics, as well as population-based surveys. In order to model AIDS-related mortality, country-specific adult HIV prevalence estimates from EPP are generated for years from the beginning of the epidemic to 2010-2015. The RUPHIVAIDS model estimates HIV incidence implied by the EPP estimates of HIV prevalence through 2010-2015, and then assumes a decline in HIV incidence of 50 percent by 2050. The graph below shows the estimated adult HIV prevalence for Malawi from 1977-2050.
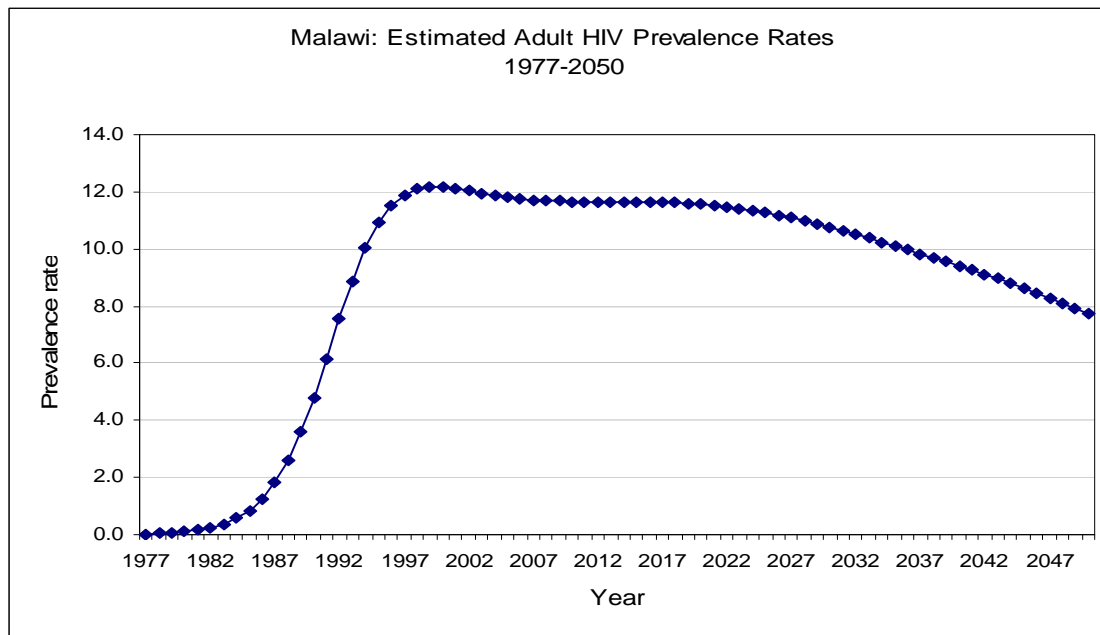
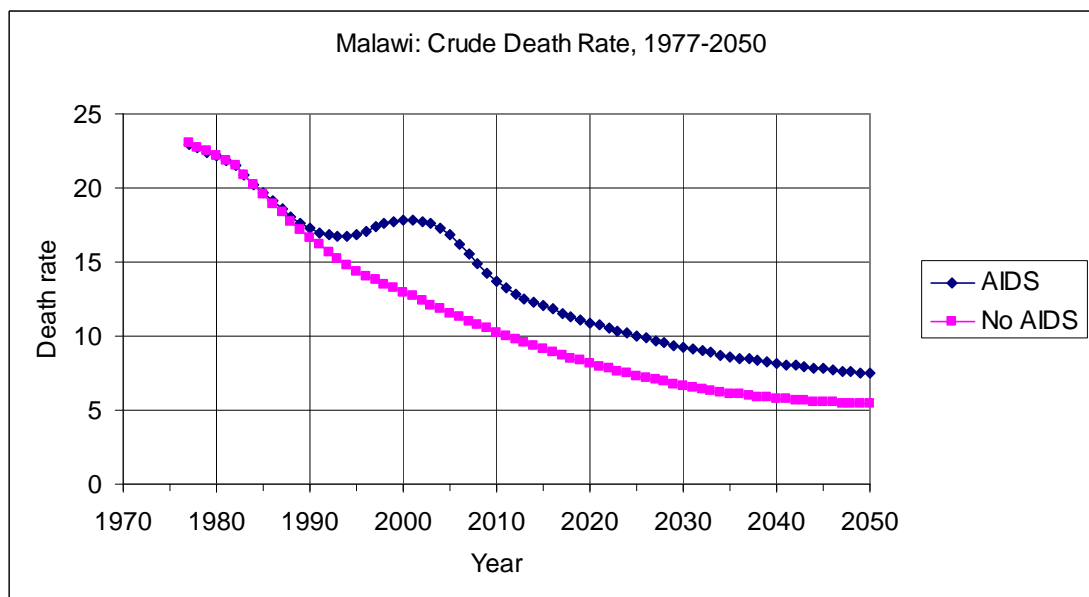**Malawi: Estimated Adult HIV Prevalence Rates
1977-2050**

Source: U.S. Census Bureau, Population Division, International Programs Center, Health Studies Branch, 2009.

In conjunction with these adult HIV prevalence estimates, RUPHIVAIDS applies assumptions from the UNAIDS Reference Group on Estimates, Modelling and Projections about the age and sex distribution of HIV incidence, sex ratios of new infections, and disease progression in both adults and children. This reference group provides the relevant technical basis for the UNAIDS/WHO global estimates and projections of HIV prevalence. The global estimates and projections represent the consensus reached at meetings held with representatives from the United Nations Population Division, U.S. Census Bureau, United Nations Children's Fund (UNICEF), WHO, and UNAIDS among others.

A hypothetical "Without-AIDS Scenario" is created using RUP to model what would have happened if a country had not been affected by the HIV/AIDS epidemic. This scenario is developed by removing estimates of AIDS mortality from observed mortality data. The resulting non-AIDS mortality is then projected into the future. The "With-AIDS Series" is then generated showing what has happened and what is projected to happen in a country as a result of AIDS mortality and its demographic consequences. The two graphs below show some mortality results under both the "Without-AIDS Scenario" and the "With-AIDS Series" for Malawi.
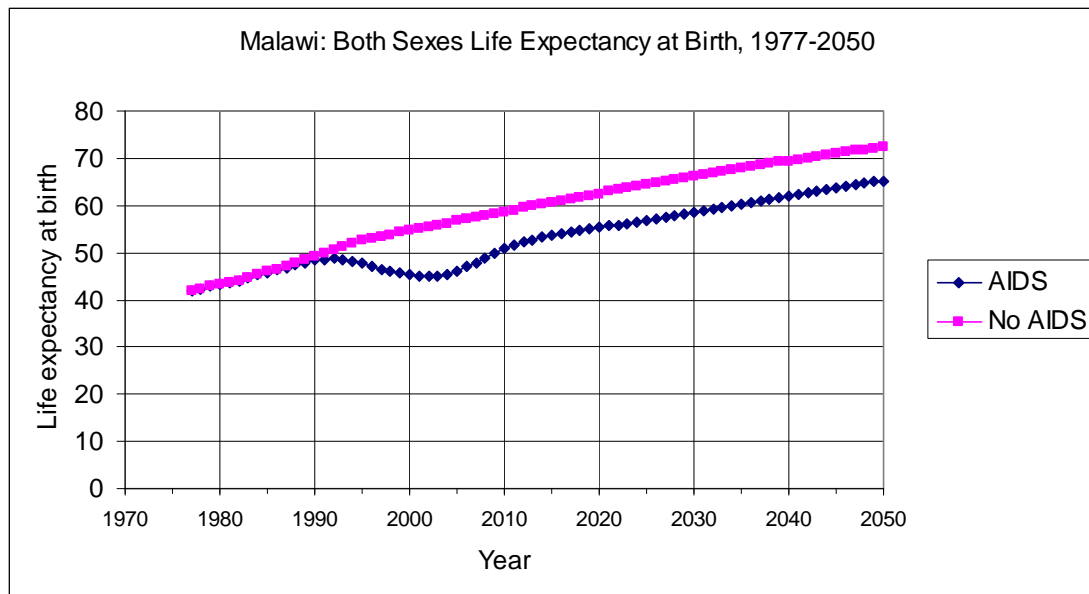
Malawi: Both Sexes Life Expectancy at Birth, 1977-2050

Malawi: Crude Death Rate, 1977-2050

The mortality estimates in the IDB include the impact of antiretroviral therapy (ART) for selected countries. The number of adults and children receiving or targeted to receive ART comes from the U.S. Office of the Global AIDS Coordinator (OGAC), WHO, and other sources. ART coverage is projected by assuming a constant yearly percent reduction in the the unmet need for ART, with the assumption that 80 percent is considered universal coverage. The yearly percent reduction is based on recent data,

projected targets, or assumed levels. If the last observed or target value is over 80 percent then that level is assumed to remain constant for the remainder of the projection.

Starting in 2008, the Census Bureau has explicitly modeled mother-to-child HIV transmission (MTCT) rates taking ART usage among pregnant women and feeding practices into account. ART is becoming more widespread through expansion of prevention of mother-to-child transmission (PMTCT) programs. The Census Bureau's current modeling entails consideration of the specific infant feeding practices and duration of breastfeeding as well as the impact of treatment to prevent transmission to the child. Information on breastfeeding practices is found in DHS surveys fielded in most of the countries as well as PMTCT program information. The proportion of women receiving single dose, dual, and triple therapy is provided by UNAIDS. Single, dual, and triple therapy are projected assuming a phase out of single-dose and dual therapy, which are replaced by the sole use of triple therapy.

Fertility: data and base estimates

As in the case of mortality, procedures for estimating fertility depend on the availability of data and on the level of details of the available data. For countries where vital registration is complete, fertility can be measured directly using classical demographic procedures. Some countries have vital registration systems which produce data that are reliable, accurate, and which capture the timing of births with relative precision. Such data tend to be released on a regular, usually annual basis, and they include births by age of mother for single years of age. In cases where births are not available, official government estimates of fertility, such as age-specific fertility rates (ASFRs) published in country statistical yearbooks or other regional statistical data bases, are accepted. Country censuses also provide fertility data, although in practice census data alone are not used to measure fertility for most countries.

Fertility for statistically less devloped countries is estimated using data from surveys, censuses and, to a lesser degree, vital registration systems. Additionally, in order to capture unusual changes in fertility levels and patterns, information from special surveys and epidemiological studies is routinely consulted.

Registered births, like registered deaths, are evaluated in order to identify potential problems. The extent and sophistication of the evaluation varies by country. Many countries have historically lacked vital registration systems that produce statistics of a quality acceptable for our demographic analysis and estimation, although during the latter half of the 20th century the systems of these countries have shown marked improvements. Before accepting registered births for projection, the Census Bureau checks available completeness ratings of the data – the United Nation's *Population and Vital Statistics Reports* provide country-specific assessments indicating if completeness is 90 percent or higher – and will evaluate births by comparing fertility levels implied by registered births with fertility estimates from other sources.

Since many countries have historically lacked reliable vital registration systems, techniques have been developed to measure fertility indirectly based on census or survey data for these countries. Some of these techniques rely on the age structure of a population from a census or the ages of children and their mothers, and an independent estimate of the level of mortality in a population. Other techniques use census and survey data from questions asked about births occurring during a fixed period of time preceding the inquiry and about the number of births a woman has ever had.

Census data that can be used in fertility analyses are available to some degree for most countries. Some fertility estimation techniques do not require the collection of census data related specifically to fertility. The age structure measured by the census can be used to estimate fertility levels through several techniques. Using the age structure of a census population, the crude birth rate is sometimes estimated by the rejuvenation technique, in which the population at the youngest ages is "reverse survived" to determine the number of births from which there are survivors. This technique is attractive because it does not require the collection of data related specifically to fertility. However, the reliability of the estimate depends on the quality of both the census data on age and the survival ratios used for the rejuvenation. Under certain circumstances, census data by age can be used to obtain not only a crude birth rate but age-specific fertility rates as well. This is done by using the own-children technique based on information on children and women by single years of age (Arriaga, 1994; Siegel and Swanson, 2004; United Nations and National Research Council, 1983).

Other techniques, such as the Rele technique, use census data by age to calculate the net reproduction rate or total fertility rate (TFR) based on the relationship of children of specified ages to the number of women in childbearing ages (Rele, 1993). The Rele technique is used in slightly different ways, depending on the quality of census data being utilized. For less developed countries with extreme underenumeration in the under-10 age groups, the Rele technique is used to establish a lower-bound fertility level. For countries whose census data appear to have relatively minimal undercount in the under-10 age group, the results of the Rele technique are given more weight and sometimes accepted as the TFR measure for a particular time point.

Many countries' censuses and surveys include questions related specifically to fertility, such as number of births in the past year and the number of children to which women have given birth. Responses to such questions can be used to evaluate fertility as well as to estimate fertility indirectly. The responses to these questions produce two measures of fertility, that, when compared, can speak to the reliability of the data. Numbers of births within the past year, per cohort of women, are used to estimate ASFRs. Similarly, the number of children ever born across a woman's reproductive life time can be used to approximate ASFRs. Two techniques utilize these comparative measures of fertility: (1) the P/F ratio technique and (2) the Arriaga technique.[14] The P/F ratio technique, developed by Brass (United Nations and National Research Council, 1983; Chapter II), is based on the average number of children ever born to women in 5-year age groups and

---

[14] Reported age-specific fertility based on vital registration data can also be used in these techniques; however, in practice, survey data and secondarily, census data, are most often used.

women's age pattern of fertility derived from births in the year preceding the census or survey. The Arriaga technique is similar to the P/F ratio technique but links data for more than one date (Arriaga, 1983 and Arriaga and Associates, 1994:233-235).
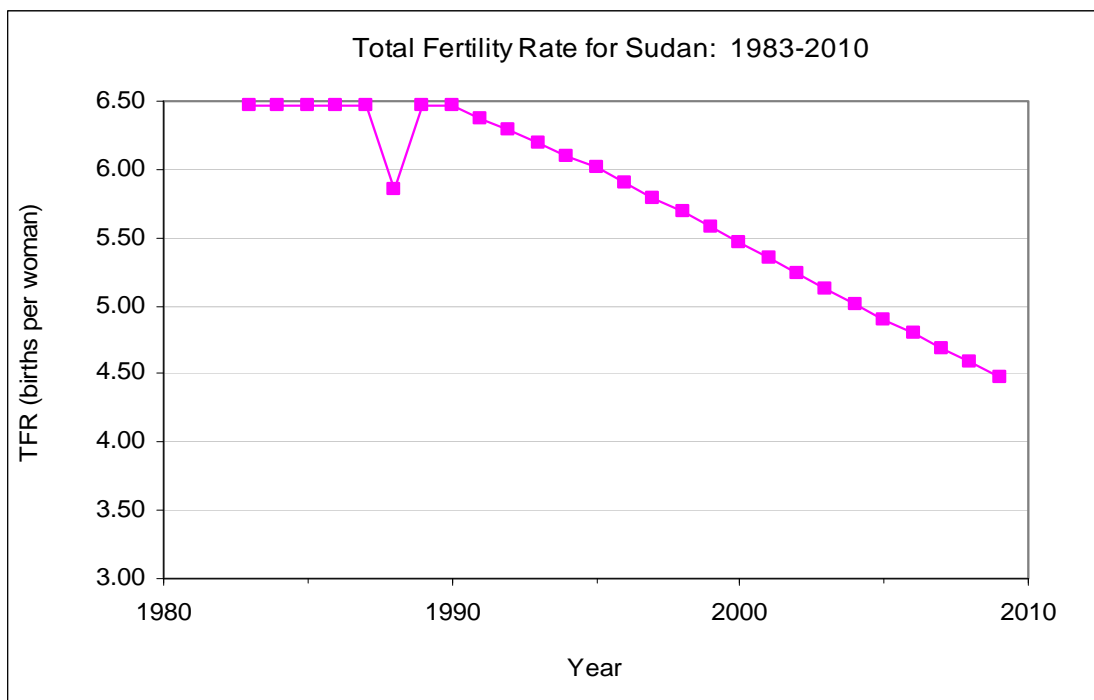
National-level surveys that provide data for fertility estimation include those undertaken as part of ORC Macro's DHS surveys, the World Fertility Surveys (WFS), the U.S. CDC Reproductive Health Surveys; Pan Arab Project for Child Development (PAPCHILD) surveys, as well as its derivatives, the Pan Arab Project for Family Health (PAPFAM) and Gulf Family Health Survey (GFHS); and United Nations Children Fund's Mutliple Indicator Cluster Surveys (MICS).   These surveys capture birth history data – that is, data on the number of births per woman and timing of births by year and age of mother. In some cases these data are of a high enough quality to use directly; however, before accepting these direct estimates for projection, they are evaluated.

In evaluating survey data, special attention is given to the coverage in surveys. Coverage in some countries can be challenging, given unstable, political and economic situations. In the African Sahel region, for example, it is difficult to interview females living in nomadic communities, and their fertility behavior is only partly understood. Similarily, women living in special areas that tend to be in conflict with the country's official government might also be under-represented in surveys.

Estimated age-sex patterns and time trends in fertility for some countries rely neither on standard demographic fertility models nor on the patterns found in similar countries due to special circumstances. Natural disasters and famines distort fertility levels and patterns, requiring additional calculations. Estimated numbers and sex composition of births may be taken from reports by various national and international organizations. Often these special fertility measures are rough estimates and they might not include complete information on the distribution of births.

Where information supporting the modeling of fertility associated with disasters is available, these data are incorporated into base fertility estimation.   Where information is not available for a particular country, information from a similar country might be used in place of data for the country under study. In some instances, simplifying assumptions are made in deciding how to adjust fertility levels and patterns. For example, a one-year dip in childbearing may be assumed for a population affected by famine, as the Census Bureau did for Sudan, which was affected by a severe famine in 1988.[15]

---

[15] Sudan fertility from update for July, 2007.

Total Fertility Rate for Sudan: 1983-2010

[Figure: Line chart titled "Total Fertility Rate for Sudan: 1983-2010". The y-axis is labeled "TFR (births per woman)" ranging from 3.00 to 6.50. The x-axis is labeled "Year" ranging from 1980 to 2010. The magenta line remains at approximately 6.50 from 1983 to about 1990 (with a downward dip to about 5.85 around 1988), then declines steadily to about 4.45 by 2010.]

Source: Based on the U.S. Census Bureau's estimates and projections update for Sudan, July, 2006.

Net international migration: data and base estimates

For some countries, data on international migration are taken directly from population registers or current official estimates prepared by the national statistical offices. For most countries, the U.S. Census Bureau usually develops migration estimates from a variety of sources, including censuses, surveys, and administrative records.

Although migration is an important component of population change for some countries, it is not generally well-recorded except in Northern and Western European countries that maintain complete and detailed population registers. Since around the first decade of the 21st century, the amount of data sources for migration has increased; and this has provided new opportunities for the U.S. Census Bureau to improve its international migration estimates. Specific data sources for migration vary widely across countries; however, some general sources consistently used by the U.S. Census Bureau for net international migration estimates include: country census data on the foreign born population, Council of Europe Statistical Yearbooks, the Migration Policy Institute (MPI) Data Hub, country arrivals and departures data, country records of visas and residence permits issued, and specialized surveys. Additionally, the following organizations publish reports on a regular basis from which the Census Bureau extracts international migration statistics: the Economic Commission for Latin America and the Caribbean (ECLAC), Eurostat, the International Labor Organization (ILO), International Organization for Migration (IOM), the Organization for Economic Cooperation and Development (OECD), the Norwegian Refugee Council, the United Nations High Commissioner for Refugees (UNHCR), the U.S. Committee for Refugees; and several

U.S. government departments, including the Department of State, Department of Homeland Security, and Department of Defense.

Some country censuses contain data on place of birth of the foreign-born population which are especially useful for estimating net migration. Net migration of the total foreign-born population over time can be measured with these data by comparing cohorts of the foreign-born across an intercensal period. With place of birth data, migration for a particular country can be gauged through the same technique. This is especially useful in trying to capture emigration levels from that country, since statistics on emigration tend to be scant. Census data on the foreign born are generally more helpful in estimating legal migration. Legal migration can also be estimated with country arrivals and departures data, if the data were processed in a way that renders them useful for statistical purposes. Similarly, if good country records are available on visas, work permits and residence permits of the foreign laborers, the Census Bureau tries to use them at least as a supplement to estimates of legal net migrants.

International refugee movements are a subset of legal cross-border migration and these kinds of movements are particularly important to demographic change in some countries. These special migration flows are incorporated by considering reported numbers of refugees documented in administrative statistics by international organizations, country sources, and media reports. In order to estimate refugee movements directly, we generally use UNHCR refugee data and subtract the refugee stock populations across two consecutive years in order to capture a net flow from one year to the next. Net flow can be more precisely tracked by examining and adjusting for data limitations. Because refugee data are collected typically in unstable situations, they often cannot capture the change in the refugee population due to births and deaths. Another issue that is commonly addressed is change in, or cessation of, refugee status. When this occurs groups of migrants might cease to show up in official accounting. Special UNHCR reports, and reports by groups such as Human Rights Watch and Amnesty International, are then relied upon to aid in quantifying the migrants. Where possible, the Census Bureau adjusts the refugee data for these population dynamics and status changes.

Refugees are largely considered as temporary migrants, since the goal of the United Nations is to repatriate them to their home country; therefore where data aren't available, the Census Bureau often assumes that most refugees will repatriate to their home country. Some refugee populations, however, are relatively long term. For these countries separate base populations are calculated and recombined to produce a total population of *de facto* residents. For example, Pakistan has hosted a large settled group of refugees and refugees' families from Afghanistan since 1979 when a large influx of Afghan refugees sought asylum there. In order to capture the population dynamics of the Afghan refugees, the U.S. Census Bureau developed a separate 1981 base population for them,[16] and estimated their population dynamics using a special set of indicators.

_____

[16] The 1981 base population for estimating and projecting Afghani refugees was developed by projecting the total 1979 refugee population to 1981, accounting for fertility and mortality. Thus, a small percentage of the 1981 "refugee" population included Afghani children, 0-2 year olds, who are not migrants. In our estimates and projections accounting process the children of Afghani refugees are placed in the category of
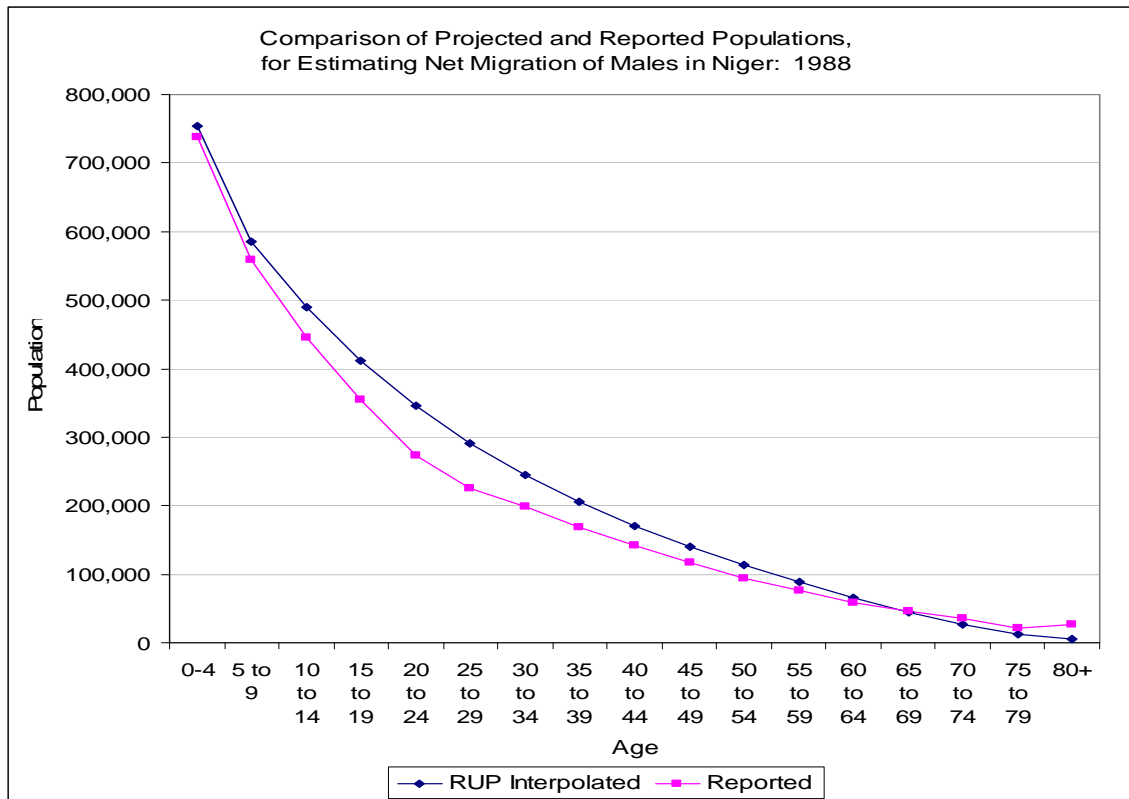
The adjustments reviewed thus far are used by the Census Bureau in estimating net migration. In some cases the Census Bureau can rely on net migration estimates produced by other agencies and organizations and will opt to insert these estimates directly into the estimates and projections process, with little or no adjustment to them. Organizations such as the Council of Europe and ECLAC, as well as country statistical agencies, sometimes publish such net migration estimates. Acceptance of net migration estimates developed by outside agencies and organizations follows a thorough evaluation of them by the Census Bureau. Evaluation entails a comparison of the organization's estimates with alternate estimates from other sources. Also, the Census Bureau might develop its own alternate set of net migration estimates as a point of comparison.

Where few or no direct estimates are available for net migration estimation and where no data are available to produce direct estimates, indirect techniques are used. Indirect methods to estimate migration vary widely by country, and it is nearly impossible to cover all the methods' nuances in the present methodology statement. Usually indirect techniques are applied through an iterative process to try to generate the best results. There are some fundamental techniques that are used, to varying degrees, to indirectly estimate migration. These include census cohort analysis and residual techniques.

Census cohort analysis entails examining population by year of birth for each sex based on data from at least two censuses. Usually the comparison is conducted by graphing 5-year and 10-year cohorts by year of birth. By graphing the population by year of birth, the comparison illustrates irregularities of the population's age structure that may have resulted from net migration. Such a comparison cannot disentangle actual demographic history from statistical errors; however, it can be useful as an initial step in examining possible migration movements.

The indirect method often used to estimate net intercensal migration for countries lacking good statistics derives migration as a residual. This residual method involves a comparison of age-sex distributions from a census or an adjusted census and an age-sex distribution resulting from a population projection that accounts for natural increase but not migration. The difference between the two populations is then attributed as net migration. The graph below illustrates the technique used to measure net migration of males in Niger between 1977 and 1988. In the graph the "Reported" line represents the reported age-sex distribution of Niger, taken from the Niger 1988 census and adjusted for coverage errors, and smoothed. The "RUP Interpolated" line represents a 1988 population projected forward from a 1977 census-based population. The gap between the two lines implies net emigration from Niger, largely between ages 5 and 59.


"refugees" since their demographic experience is more similar to their Afghani refugee parents, even though Pakistani citizenship rules allow for the acquisition of Pakistani citizenship by birth.

Comparison of Projected and Reported Populations, for Estimating Net Migration of Males in Niger: 1988

Population

800,000
700,000
600,000
500,000
400,000
300,000
200,000
100,000
0

0-4 | 5 to 9 | 10 to 14 | 15 to 19 | 20 to 24 | 25 to 29 | 30 to 34 | 35 to 39 | 40 to 44 | 45 to 49 | 50 to 54 | 55 to 59 | 60 to 64 | 65 to 69 | 70 to 74 | 75 to 79 | 80+

Age

—◆— RUP Interpolated —■— Reported

Source: Based on the U.S. Census Bureau's estimates and projections update for Niger, December, 2008.

For some estimates, the age-sex distribution implied by the comparison is used in combination with official statistics from country governments. This technique of combining data was used for estimating Canada's net migration. Specifically official statistics on migration totals were accepted and then disaggregated by age and sex using the distribution implied by the residual procedure.[17] Below is an example of the age-sex distribution for females implied by a comparison of a 1996 projection with an adjusted 1996 census population for Canada.

---

[17] This technique was used for the estimates and projections update of Canada in March, 2008.

Comparison of Projected and Reported Populations, for Estimating Net Migration of Females in Canada: 1996
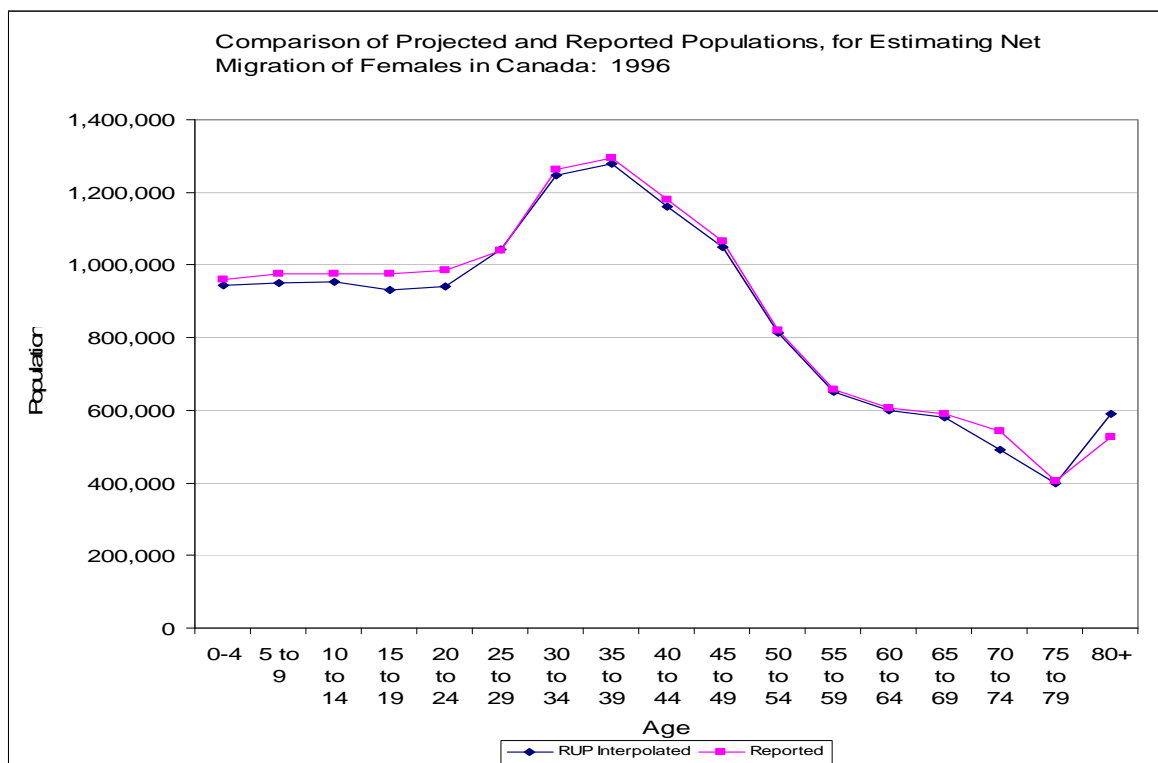
Source: Based on the U.S. Census Bureau's estimates and projections update for Canada, March, 2008.

The residual technique is used, as needed, in migration estimation for most countries, sometimes in combination with data on international migration available from population registers, censuses, surveys, and administrative records.

Indirect techniques are often used in cases where a substantial proportion of migrants are unauthorized, since unauthorized migrants are often not captured adequately. Additionally, sometimes the Census Bureau takes information from media or special migration news reports and studies. Destination country records on immigrants, too, can provide at least a partial "mirror" of emigration for some countries. For many countries, a combination of indirect techniques and data and direct data sources are used. Because of the uncertainty involved in estimating unauthorized migration, broad and sometimes extensive assumptions are necessarily made. In analyzing data on unauthorized migration and formulating assumptions, special attention is often given to the principal countries of origin and destination.

Projecting Population and Components of Change

Once levels of mortality, fertility, and net international migration have been estimated, each component must be projected into the future. Although the procedure for doing this

is often mechanical, care must be taken in determining projected levels, trends, and patterns by age. Not only must the assumptions be appropriate for the particular country in question, but consistent assumptions must be made across countries

Mortality

Some assumptions about mortality trajectories are consistent across countries and regions. In general, mortality is expected to continue to decline in most countries, as economic development occurs and health care improves. A particular exception relates to the impact that AIDS will have on the mortality of some countries, where mortality levels in the next decade are expected to increase. While there is no single "right" way to make assumptions about the future, the Census Bureau relies heavily on extrapolation of past trends in indicators, coupled with validation checks against published estimates of determinants and correlates in preparing assumptions about future mortality trends.[18] S-shaped logistic functions are typically used to model the transition from relatively high mortality to relatively low mortality.

In order to project future mortality levels, the Census Bureau generally fits a logistic curve to one or more estimates of life expectancy at birth. The results of the logistic projection are carefully scrutinized, however, to ensure that they yield an acceptable projected level for the given individual country's circumstances.[19]

More often than not, the Census Bureau uses a variant of the basic logistic to project e(0) that assumes the same slope for each country. This variant, developed at the Census Bureau in the late 1990s by fitting the logit transformation of e(0) for a number of countries and denoted as the "fixed-slope" logistic, uses slope values of 0.0258 for males and 0.0271 for females.

---

[18] Correlates of mortality include childhood immunization coverage, for example.

[19] In using the logistic function to project life expectancy at birth (e(0)), the Census Bureau utilizes sex-specific upper and lower limits, called "asymptotes," to control the values projected by sex.   The formulation of the standard logistic function expresses predicted e(0) as a function of defined lower and upper asymptotes and the estimated slope and intercept of an ordinary least squares regression line fitted to the logits of estimated e(0)s. Logit transformations are calculated as:

$$\text{logit}(e(0)t) = \ln\left[\frac{(e(0) - LB)/(UB - LB)}{1 - (e(0) - LB)/(UB - LB)}\right]$$

where UB is the upper asymptote for the logistic function, LB is the lower asymptote for the logistic function, and t is time in years.

Predicted values based on the fitted, linear relationship, $\text{logit}(e(0)t) = \hat{a} + \hat{b}t$, are then converted to predicted e(0)s:

$$\hat{e}(0) = \left(\frac{e^{a+bt}}{1 + e^{a+bt}}\right) * \left(UB - LB\right) + LB$$

where a is the intercept of a line fitted to the logit transformations and b is the slope of a line fitted to the logit transformations.

The following graphs compare the results of projecting life expectancy at birth along fitted and fixed curves for males and females. In the fitted variant, a logistic curve is fitted to observed life expectancy at birth for at least two reference points in order to predict the slope of future decline through 2050. In the fixed variant, observed life expectancy at birth doesn't change the slope; however, the observed data influence, as they do in the fitted slope function, the overall level of life expectancy values across the time period under study.

In the example shown, observed life expectancies at birth for males were 46.95 for 1995.7 and 48.43 for 1998.2. Using these observed values to project life expectancy along a fixed slope generates a life expectancy at birth of 65.91 for 2050. Fitting a logistic to these observed values instead results in a life expectancy at birth value of 74.12 for 2050.
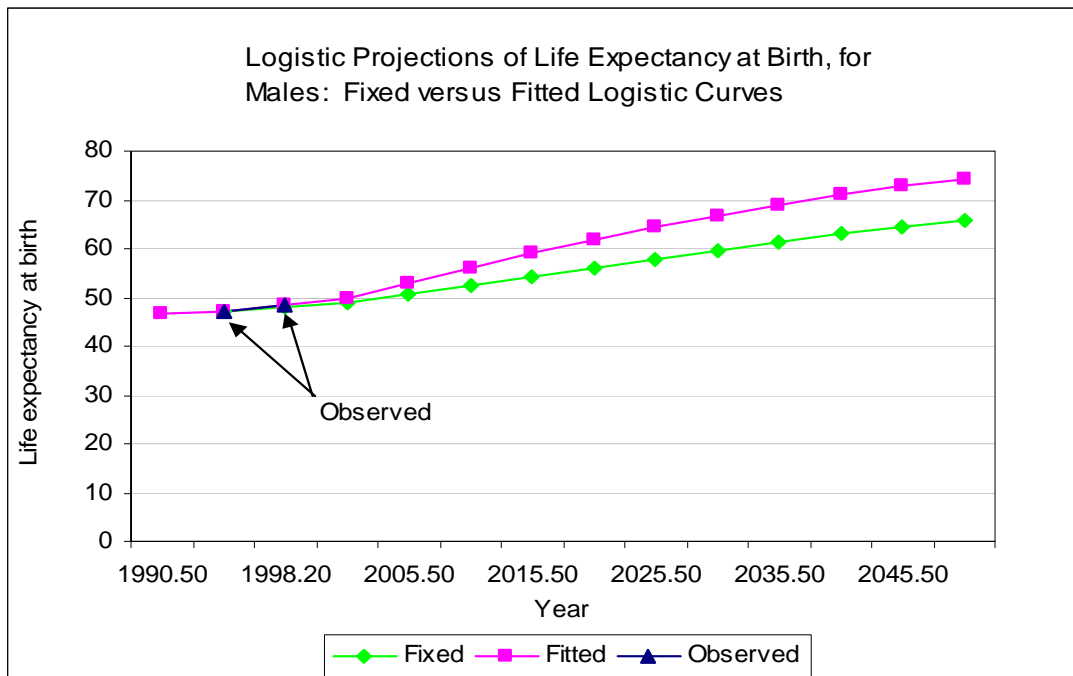
Source: U.S. Census Bureau, Population Division, International Programs Center, Population Studies Branch, 2009.

The results of logistic projections of e(0) are evaluated in light of recent socioeconomic trends, social policies, and public health and program coverage. Once mortality has been tentatively projected for each country according to its particular circumstances, the projections are compared with projected values for other countries in the same region.

After assumptions about future levels of components are developed, age patterns of mortality for each of the projected values are developed, since these patterns tend to vary

as overall levels change. For each level of projected life expectancy at birth, a set of central death rates is estimated using an iterative interpolation process. The interpolation is logarithmic and uses a set of central death rates for the last input year—an "ultimate" set of rates with very low mortality. Life tables constructed with the interpolated rates correspond to the life expectancies at birth projected previously.

Finally, distinctive mortality assumptions are made for selected countries because of the death risk due to AIDS. Using methodology that takes into account the effect of AIDS, country projections are prepared that assess its impact on future populations in countries where the HIV prevalence is significant.

Fertility

As in the case for mortality, some assumptions about the fertility trajectories are consistent across countries and regions. An expected increase in contraceptive prevalence is implicit in the assumptions about future fertility declines for many countries. For some countries, future fertility levels are projected to experience only minor change, either slight decreases or slight increases. While there is no single "right" way to make assumptions about the future, the Census Bureau relies heavily on extrapolation of past trends in indicators, coupled with validation checks against published estimates of determinants and correlates in preparing assumptions about future fertility trends. Logistic functions are typically used to model the transition from relatively high fertility to relatively low fertility.

In order to project future fertility levels, the Census Bureau generally fits a logistic curve to one or more TFR estimates. If estimates of TFR are available for more than one date in the past and the TFR is not already below 1.7, a logistic function is fitted to these data. The results of the logistic projection are carefully scrutinized, however, to ensure that they yield an acceptable projected level for the given individual country's circumstances. In some instances, no data on past trends in fertility are available for fitting a logistic curve. In that case, the past experience of neighboring or similar countries serves as a guide for fitting the likely pace of future change.

A logistic function is typically used to project TFRs to 2050 with lower and upper limits depending on the current level of fertility in a country. There are some commonalities among regions, however. Regions which tend to be transitioning from higher to lower fertility have high TFR limits of up to an average of 9 births per woman and a lower limit for 2050 of 2.[20] An example of TFR projected along a logistic curve with an upper asymptote of 7 and a lower asymptote of 2 is shown here.

---

[20]The functional form of the standard logistic curve used to project TFR is:

$$\hat{TFR}_t = \left(\frac{e^{a+bt}}{1 + e^{a+bt}} \right) * (UB - LB) + LB$$
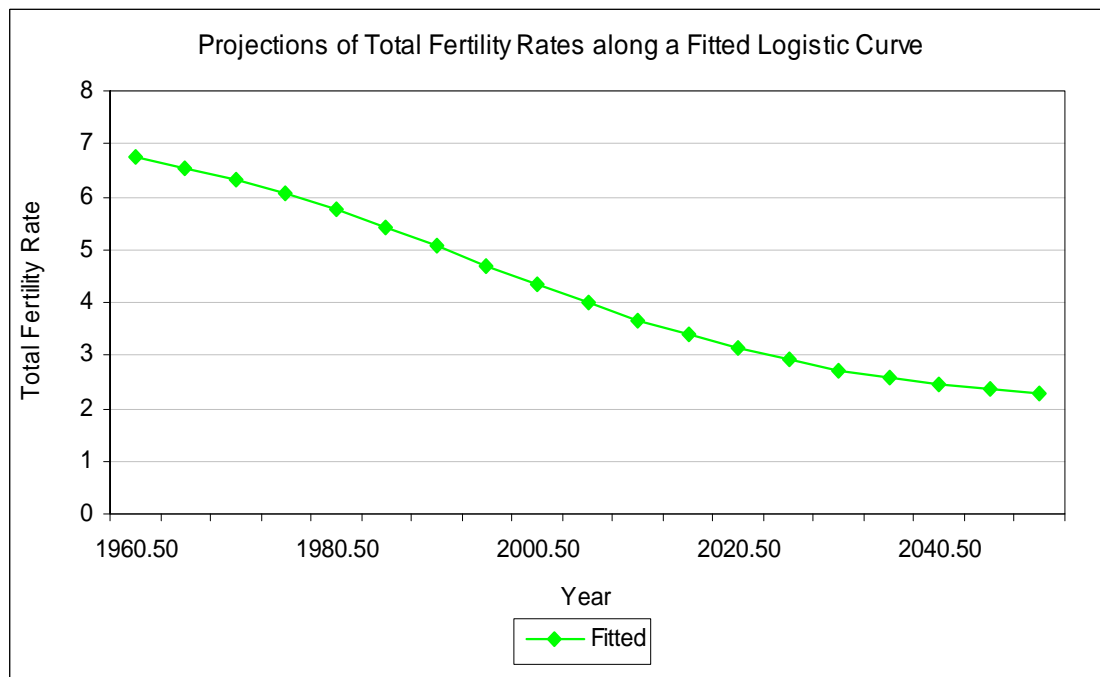
where

Projections of Total Fertility Rates along a Fitted Logistic Curve

Source: U.S. Census Bureau, Population Division, International Programs Center, Population Studies Branch, 2009.

The results of logistic projections are evaluated in light of recent socioeconomic trends, social policies, public health and program coverage, and the proximate determinants of fertility. Trends in women's educational attainment and labor force participation, for example, are examined. Family planning program efforts are also reviewed, as are age at marriage and the proportion of women using contraception. Recent data on the current use of family planning methods are gathered primarily by surveys such as the DHS, the family health and contraceptive prevalence surveys of the CDC, and the various survey programs in Europe and the Middle East. Once fertility has been tentatively projected for

---

UB	is the upper asymptote for the logistic function
LB	is the lower asymptote for the logistic function
b	is the slope of a line fitted to the logit transformations of observed TFRs
a	is the intercept of a line fitted to the logit transformations of observed TFRs
t	is time in years

and the logit of each observed TFR is

$$\ln\left[\frac{(TFR - LB)/(UB - LB)}{1 - (TFR - LB)/(UB - LB)}\right]$$

each country according to its particular circumstances, the projections are compared with projected values for other countries in the same region.

After assumptions about future levels of TFR are developed, age patterns of fertility for each of the projected values are developed, since these patterns tend to vary as overall levels change. For each level of projected TFR, a set of ASFRs is estimated using an iterative interpolation process. The interpolation is logarithmic and uses a set of ASFRs for the last input year and an "ultimate" set of rates with very low fertility. The ultimate set of rates was derived at the International Programs Center using empirical data for countries with low fertility.

Net international migration

Assumptions about future net international migration are generally much more speculative than assumptions about fertility and mortality. International migration may occur as a result of changing economic conditions, political unrest, persecutions, famines, and other extreme conditions in the countries of origin as well as destination.

Due to the unpredictability of conditions such as crop failure, emerging violence, and war, migration forecasts are subject to large errors. If migration is known to have a negligible impact on a country's current growth rate, future migration is often assumed to be nil. If a country's migration is known to be significant, the estimated number of migrants during the past is frequently held constant in projecting to the near future. The age and sex composition of international migrants depends on the situation in each country. If information is not available, model patterns of migrant age and sex composition are sometimes used.
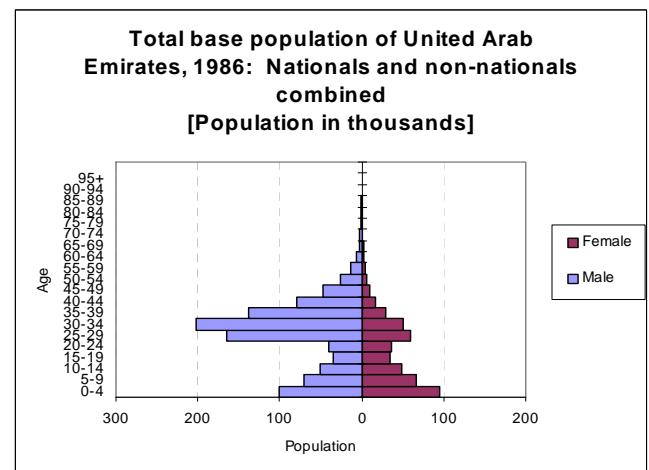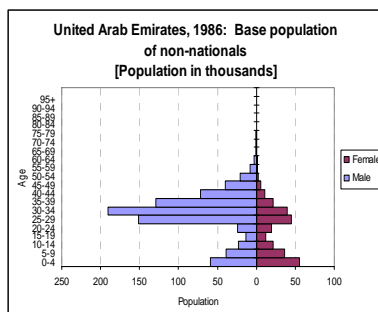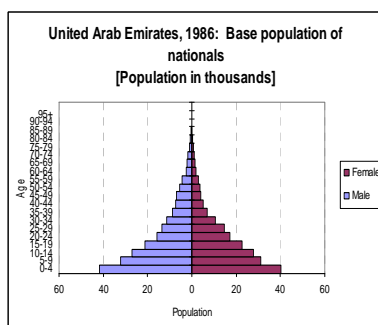
Population estimates and projections for countries are revised on a flow basis. Therefore, migration streams across all countries may not balance to zero, and net migration for the world may be positive or negative.

For some countries the number of net international migrants and their movements from year to year are substantial enough, predictable enough, and are represented by enough data of acceptable quality in order to allow for separate cohort component projections of them. Such is the case in many Persian Gulf countries where large portions of individual country populations consist of labor migrants who enter the countries to work on a regular basis. For these countries, separate cohort-component projections for the national population and the non-national population of migrants are made and then recombined to produce a composite set of cohort-component projections for the entire country. In other words, instead of simply adding (or subtracting) net international migrants by age and sex to U.A.E.'s national population; the international migrant population, itself, is estimated and projected to 2050 in a more complex and precise manner, using the cohort-component method.

Below is an example of the U.A.E. cohort component projections of nationals and non-national migrants and the resulting composite cohort-component projection representing the *de facto* population of U.A.E.

In the first set of population charts, representing the 1986 base population, the base year age-sex composition for the national and non-national populations of the U.A.E. are shown on the left; the national population is on the top left and the non-national population on the bottom left. The atypical structure for the non-national population, dominated by males of labor force age, influences the age-sex composition profile of the total U.A.E. population of nationals and non-nationals combined, shown at the right.
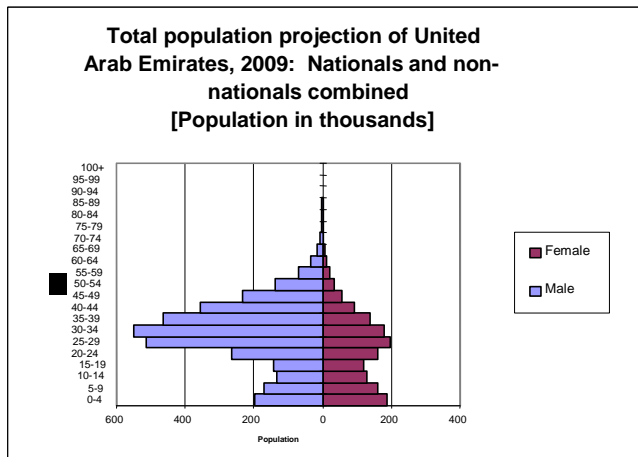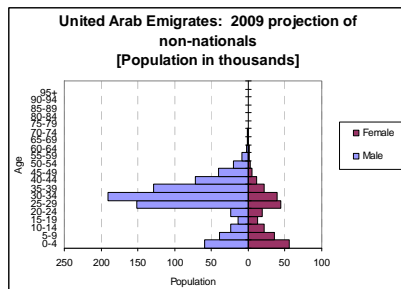
1986 Base population

United Arab Emirates, 1986: Base population of nationals
[Population in thousands]

Total base population of United Arab Emirates, 1986: Nationals and non-nationals combined
[Population in thousands]

United Arab Emirates, 1986: Base population of non-nationals
[Population in thousands]
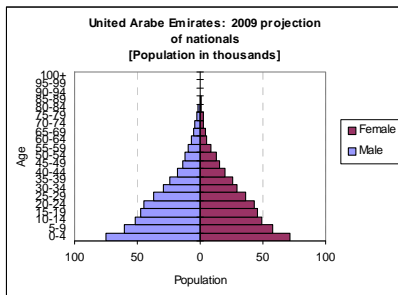
Source: Based on the U.S. Census Bureau's estimates and projections update for U.A.E., July, 2007.

Now here we see, the separate 2009 cohort-component projections of nationals and non-nationals, shown at the left. The composition profile on the right again reflects the male-dominated non-national age-sex structure.
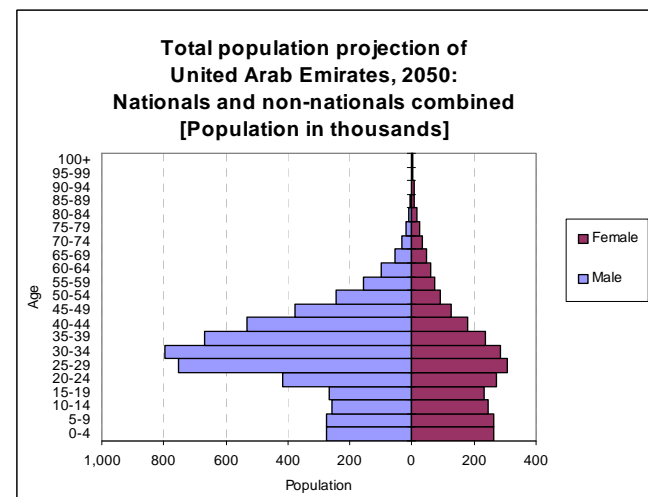
2009 projection

United Arabe Emirates: 2009 projection of nationals
[Population in thousands]

Total population projection of United Arab Emirates, 2009: Nationals and non-nationals combined
[Population in thousands]

United Arab Emigrates: 2009 projection of non-nationals
[Population in thousands]
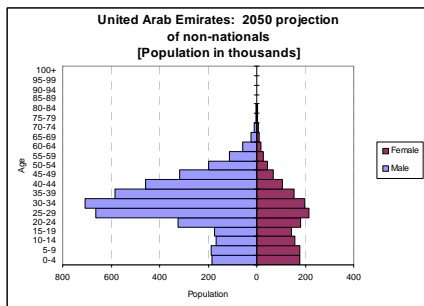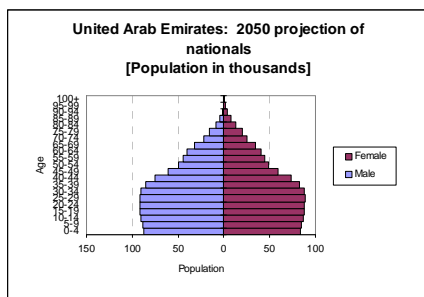
Source: Based on the U.S. Census Bureau's estimates and projections update for U.A.E., July, 2007.

Again, the separate cohort-component projections of nationals and non-nationals to 2050 reveal very distinct patterns of nationals and non-nationals, as shown at the left. The composition profile on the right sustains in large measure the influence of the male-dominated non-national age-sex structure.

2050 projection

United Arab Emirates: 2050 projection of nationals [Population in thousands]

United Arab Emirates: 2050 projection of non-nationals [Population in thousands]

Total population projection of United Arab Emirates, 2050: Nationals and non-nationals combined [Population in thousands]

Source: Based on the U.S. Census Bureau's estimates and projections update for U.A.E., July, 2007.
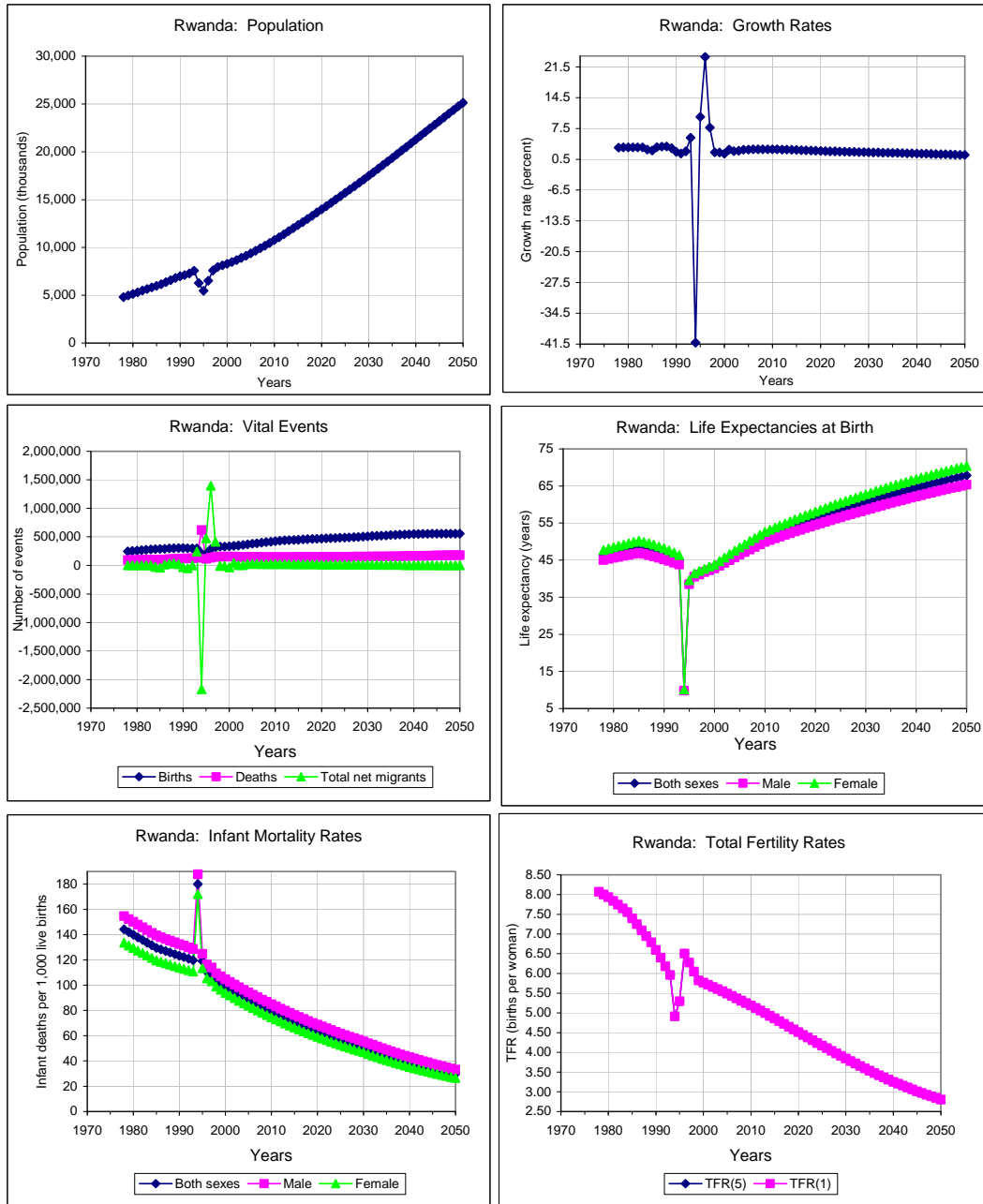
Estimating demographic shocks

The Census Bureau estimates and projects the impact of demographic shocks on population and components of change. Examples of demographic shocks include war, famine, other natural disasters, and severe and sudden economic downturns. Data sources for the timing of and deaths resulting from such events, include the Emergency

Events Database (EMDAT), news articles, deaths tallies produced by the WHO, special surveys, and epidemiological studies conducted by various government agencies and research groups (including the CDC's National Center for Health Statistics, etc.). Fertility impacts are more difficult to quantify, but some evidence is available from post-event census and survey data and from news articles. International migration flows attributed to demographic shocks are documented by the UNHCR and other organizations listed earlier on source of migration information, and news articles.

The Census Bureau translates the short-term properties of demographic shocks into estimates and projections by specifying the time period (half-year, year, or range of years) during which the event occurred as part of a country's projection assumptions. RUP enables the Census Bureau to limit the immediate impact of an event on the components of change to as little as a half-year period. The after-effects and gradual return of population size and patterns to pre-shock level are also integrated into assumptions. For example, in 1994 the population of Rwanda suffered a genocide during April, May, and June. The Census Bureau was able to estimate the death toll and immediate change in population size by isolating the event to the first half of 1994. Subsequent changes in other population components, the gradual reduction in the mortality rate, and the replenishment of the population over time was then modeled in the projections through subsequent years to 2050. The following charts illustrate the Census Bureau's measurement of the impact of genocide in Rwanda in 1994 on projected population and other demographic indicators.[21]

---

[21] From estimates and projections update August, 2006.

Rwanda: Population

Rwanda: Growth Rates

Rwanda: Vital Events

Rwanda: Life Expectancies at Birth

Rwanda: Infant Mortality Rates

Rwanda: Total Fertility Rates

Source: Based on the U.S. Census Bureau's estimates and projections update for Rwanda, August, 2006.

Pre-base Year Data in the IDB

Due to the differing nature of the base data for each country, there is no standard starting date for each country's cohort-component projection. The projection period may start only a few years ago, when the base information was current to that date. On the other hand, the projection period may start several decades ago although information from a later date for one or more of the variables may have been taken into account when deriving parameter estimates for the early years of a projection. "New" information for a country whose projection base year is in the 1970s may pertain to the 1980s or 1990s, superseding the 1970s data available for a previous round. Thus, the total population in the revised projection may change for any year in the past. Components also can change in this way. Population age structure and components of change are available in the IDB for each country from the base year onward.

Total midyear population estimates and projections in the IDB are available for 1950 onward for all current countries whether or not they existed in 1950. The data and methods through which the 1950 populations were developed vary depending on the availability of data per country. For many less developed countries, modern censuses were not conducted until the 1960s or 1970s, and reliable population data therefore are not available. Midyear population for the 1950s in the IDB has been estimated using census data for later years and assumptions about population growth rates for countries. Total midyear populations for each year between 1950 and the base year of the cohort-component population estimates and projections are revised periodically; specifically, when the base year population of the cohort-component projections is revised.

Total midyear populations pertaining to the base year from which the cohort-component projections are developed are changed periodically; specifically when new data, or a reinterpretation of older data, suggests they should be changed. Base population totals often change when the age and sex composition of the population is revised. The age and sex composition of the base population is revised oftentimes in order to render it consistent with new assumptions about fertility, mortality, and the age-sex distribution and size of a more recent census.

Sources and Accuracy

The Census Bureau prepares national estimates and projections for all countries using census and survey data, vital statistics, administrative statistics from the countries, and information from multinational organizations that collect and publish data for these countries. The Census Bureau's evaluation process focuses on internal and temporal consistency of data. Information on statistical measures of uncertainty (e.g., standard errors) are considered in the evaluation when such data are available.

References

Attama, Sabine, Michka Seroussi, Alichina Idrissa Kourguéni, Harouna Koché et Bernard Barrèm. 1998. *Enquête Démographique et de Santé, Niger 1998.* Calverton, Maryland, U.S.A.: Care International and Niger et Macro International Inc.

Arriaga, Eduardo E. 1983. *Estimating Fertility from Data on Children Ever Born, by Age of Mother.* U.S. Bureau of the Census, International Research Document No. 11. ISP-RD-11. Washington, D.C.: U.S. Government Printing Office.

Arriaga, Eduardo E. and Associates. 1994. *Population Analysis with Microcomputers.* U.S. Census Bureau, International Programs Center: Washington, D.C.

Bachi, Roberto. 1951. "The Tendency of Round Off Age Returns: Measurement and Corrections," in *Bulletin of the International Statistical Institute*, Vol. 33, No. 4, pp. 195-222.

___. 1953. "Measurement of the Tendency of Round Off Age Returns," in *Bulletin of the International Statistical Institute*, Vol. 34, No. 3, pp. 129-137.

Bhat, Mari PN. 2001. "Completeness of India's sample registration system: an assessment using the general growth balance method." *Population Studies* (Camb). Delhi, India: Institute of Economic Growth.

Bogue, Donald J., Eduardo E. Arriaga, and Douglas L. Anderton. 1993. *Readings In Population Research Methodology.* Chicago, I.L.: Social Development Center.

Bongaarts, John. 1984. "A simple method for estimating contraceptive prevalence required to reach a fertility target." *Studies in Family Planning*, Vol. 15, No. 4 (Jul-Aug., 1984), pp. 184-190.

Brass, William, et al. 1968. Chapter III in *The Demography of Tropical Africa*. Princeton, N.J.: Princeton University Press.

Brass, William. 1975. *Methods for Estimating Fertility and Mortality from Limited and Defective Data*, International Program of Laboratories for Population Statistics, The Carolina Population Center, University of North Carolina at Chapel Hill.

___. 1981. "The use of Gompertz relational model to estimate fertility," in *Proceedings of the International Population Conference, Manila*, International Union for the Scientific Study of Population, Vol. 3, pp. 345-362. Liège.

Cho, Lee-Jay. 1973. "The own-children approach to fertility estimation: an elaboration," in *Proceedings of the International Population Conference, Liège 1973*, International Union for the Scientific Study of Population. Liège.

Coale. Ansley J. and Paul Demeny. 1966. *Regional Model Life Tables and Stable Populations*. Princeton, N.J.: Princeton University Press.

Coale, Ansley J. and Guang Guo. 1989. "Revised Regional Model Life Tables at Very Low Levels of Mortality." *Population Index*, Vol. 55, No. 4, pp. 613-643.

___. 1990. "Revised Regional Model Life Tables at Very Low Levels of Mortality." [Full set of appendix tables and an additional reference]. *Population Index*, Vol. 56, No.1, pp. 26-41.

Davis, Kingsley and Judith Blake. 1956. "Social structure and fertility: An analytic framework." *Economic and Cultural Change*, Vol. 4. No. 2, pp. 211-235.

Ghys P., N. Walker, W. McFarland, et al. 2008. "Improved data, methods and tools for the 2007 HIV and AIDS estimates and projections." *Sexually Transmitted Infections*, Vol. 84 (Supplement I), pp. i1-i4.

Hill, K. 1987. "Estimating census and death registration completeness." *Asian Pacific Population Forum*. May; Vol. 1, No. 3, pp. 8-13 and pp. 23-4.

Hobbs, Frank. 2004. "Age and Sex Composition." Chapter 7 in *The Methods and Materials of Demography,* Second Edition.   Edited by Jacob S. Siegel and David A. Swanson. New York, N.Y.: El Sevier Academic Press.

Institut National de la Statistique (INS) et Macro International Inc. 2007. *Enquête Démographique et de Santé et à Indicateurs Multiples du Niger 2006*. Calverton, Maryland, USA : INS et Macro International Inc.

Johnson, Peter. 2006. Extension of the Population to Age 100+ Using the Pop100 Workbook. Unpublished paper. Washington, D.C.: U.S. Census Bureau, International Programs Center.

___. 2006. Revised Methods for Extending Mortality to 100+ Using the LTEXTNew and LTSPLTNew Workbooks. Unpublished paper. Washington, D.C.: U.S. Census Bureau, International Programs Center.

Joint United Nations Programme on HIV/AIDS. 2007. "Financial Resources Required to Achieve Universal Access to HIV Prevention, Treatment, Care and Support." <http://data.unaids.org/pub/Report/2007/20070925_advocacy_grne2_en.pdf>. Geneva, Switzerland.

Joint United Nations Programme on HIV/AIDS/World Health Organization. 2007. The UNAIDS/WHO Estimation and Projection Package (EPP 2007). Version 3.0.

___. 2009. Estimating National Adult Prevalence of HIV−1 In Generalized Epidemics. Draft Manual. <http://data.unaids.org/pub/Manual/2009/20090407_epp_genepi_2009_draft_en.pdf>. Geneva, Switzerland.

Kourgueni, Idrissa A., Garba, Bassirou, and Barrere, Bernard. 1993. Niger: Enquete Demographique et de Sante 1992. Columbia, M.D.: Macro International.

Myers, Robert. 1940. "Errors and Bias in the Reporting of Ages in the Census Data," in *Transactions of the Actuarial Society of America*, Vol. 41, Part 2, pp. 411-415.

Popoff, Carol and D.H. Judson. 2004. "Some Methods of Estimation for Statistically Underdeveloped Areas." Chapter 22 in *The Methods and Materials of Demography,* Second Edition. Edited by Jacob S. Siegel and David A. Swanson. New York, N.Y.: El Sevier Academic Press.

Rele, J.R. 1967. *Fertility Analysis Through Extension of Stable Population Concepts*, Population Monograph Series, No. 2, Institute of International Studies, University of California, Berkeley.

___. 1993. "The Rele Procedure for Indirect Estimation of Fertility." Chapter 9 in *Readings In Population Research Methodology.* Edited by Donald J. Bogue, Eduardo E. Arriaga, and Douglas L. Anderton. Chicago, I.L.: Social Development Center.

Siegel, Jacob and David. A. Swanson. 2004. *The Methods and Materials of Demography,* Second Edition.   New York, N.Y.: El Sevier Academic Press.

Stover, John. 1998. "Revising the Proximate Determinants of Fertility Framework: What have we learned in the Past 20 Years?" *Studies in Family Planning*. Vol. 29, No. 3.

___, P. Johnson, B Zaba, et al. 2008. The Spectrum projection package: improvements in estimating mortality, ART needs, PMTCT impact and uncertainty bounds. *Sexually Transmitted Infections*, Vol. 84 (Supplement I), pp. i24-i30.

Thatcher, A.R., V. Kannisto, and J.W. Vaupel. 1998. *The Force of Mortality at Ages 80 to 120*. Monographs on Population Aging, 5. Odense, Denmark: Odense University Press.

United Arab Emirates, Ministry of the Economy. (?). Statistic Abstract 2007. <http://www.economy.ae/English/EconomicAndStatisticReports/StatisticReports/StatisticAbstract/Pages/sa2007.aspx >

United Kingdom Border Agency. Vol. I. The British Nationality Act caseworking instructions. Chapter 14, Annex J: Pakistani Citizenship Law. [Accessed 3/20/09] http://www.ukba.homeoffice.gov.uk/sitecontent/documents/policyandlaw/nationalityinstructions/nichapter14/

United Nations. 1952. "Accuracy tests for census age distributions tabulated in five-year and ten-year groups," in *Population Bulletin*, no. 2, ST/SOA/Ser.N/2, pp. 59-79. New York, N.Y.

___. 1955. *Methods of Appraisal of Quality of Basic Data for Population Estimates.* Manual II, Series A, Population Studies No. 23, pp. 42-43. New York.

___. 2000. *United Nations Demographic Yearbook: Historical Supplement 1948-1997.* Geneva, Switzerland.

___. 2003. *World Population Prospects: The 2002 Revision*. "File 5: Life Expectancy at Exact Age, e(x), for Both Sexes Combined, by Major Area, Region and Country, 1990-2050." Data file POP/DB/WPP/Rev.2002/5/F5. New York: UN Population Division.

United Nations and National Research Council. 1983. *Manual X. Indirect Techniques for Demographic Estimation.* New York, N.Y.: United Nations.

U.S. Bureau of the Census. 1971. *The Methods and Materials of Demography*, by Henry Shryock, Jacob Siegel and Associates, Vol. 1, pp. 205-211. Washington, D.C.: U.S. Government Printing Office.

___. 1983. "Estimating fertility from data on children ever born, by age of mother," by Eduardo E. Arriaga, International Research Document No. 11. Washington D.C.

U.S. Census Bureau. 2001. *Population Estimates and Projections Guidelines.* Washington, D.C.: Population Division, International Programs Center. Unpubublished.

___. 2004. Unpublished life table values from RUP projections for countries modeled up to age 100+ in the International Data Base, September 30, 2004 update. Washington, DC: Population Division, International Programs Center.

___. 2005. Unpublished work tables for estimating Iran's mortality. Washington, D.C.: Population Division, International Programs Center.

___. 2009. Unpublished work tables for estimating and projecting populations for the International Data Base, various years. Washington, D.C.: Population Division, International Programs Center.

___. 2009. Unpublished work tables for estimating and projecting HIV/AIDS prevalence and mortality levels with AIDS, various years. Washington, D.C.: Population Division, International Programs Center, Health Studies Branch.

___. 2009. Unpublished work tables for estimating and projecting HIV/AIDS prevelance, various years. Washington, D.C.: Population Division, International Programs Center, Population Studies Branch.

World Health Organization. 2008. "Towards Universal Access: Scaling up Priority HIV/AIDS Interventions in the Health Sector, Progress Report 2008." <http://www.who.int/hiv/mediacentre/2008progressreport/en/index.html>. Geneva, Switzerland.